

Everything Was Good!

The Influence of Sentiment and Product Category on Aspect Choice in German Customer Reviews

Amelie I. Metzmacher^a, Verena Heinrichs^a, Björn Falk^a, Robert H. Schmitt^a

^aLaboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Aachen, Germany
E-mail: A.Metzmacher@wzl.rwth-aachen.de, V.Heinrichs@wzl.rwth-aachen.de, B.Falk@wzl.rwth-aachen.de,
R.Schmitt@wzl.rwth-aachen.de

Abstract— Studies dealing with the use of pronouns in customer reviews in social media against the background of sentiment analysis do not consider the role attributed to the pronoun within an evaluative sentence. The following paper addresses the question of when and why customers choose pronouns instead of nouns or proper nouns to function as aspects within an evaluative sentence. To this end, two influencing factors on aspect choice have been investigated: 1) The sentiment of the evaluative sentence (positive, neutral, negative); 2) The category of the evaluated aspect (product, service). The results reveal that, particularly within positive statements, there exist differences between customer reviews evaluating products to those evaluating services.

Keywords-NLP; sentiment analysis; social media.

I. INTRODUCTION

With respect to the growing demand for individualized products and services, understanding the Voice of the Customer (VoC) becomes increasingly important for companies. To better survive competition and fulfill customer demands, it is crucial for companies to develop customer-oriented solutions based on precise knowledge about customer requirements [1].

Knowledge about customer requirements can be captured by studies, interviews, ethnographic research, customer visit teams, customer brainstorming, or lead user analysis [2]. However, these methods are most often time consuming and costly while suffering from low response rates.

Social media provide an alternative to gain knowledge about customer requirements. Information relating to customer requirements derived from social media is more authentic as customers in discussion forums usually uninhibitedly and immediately provide their opinion about products and services [3]. Moreover, the amount of data available online is steadily increasing.

In consideration of the steadily increasing number of available data, it is not possible to extract relevant information from customer reviews in social media manually. Thus, it is necessary to develop a method to automatically extract only the beneficial information from customer reviews, i.e., information about the perceived quality of products and services [4].

Aiming at a supervised machine-learning based solution, an annotation study was conducted to obtain an annotated corpus of German customer reviews from social media. Although there already exist such corpora, e.g., cf. [5], these corpora only include reviews relating to material products.

Within our approach, we focus on both, reviews relating to products and to services. The consideration of products and services is particularly helpful regarding the application for companies. Often reviews about products, as well as services need to be taken into account for customer-oriented product development and design. An example for this is a company producing domestic appliances while at the same time offering services for maintenance, repair and operations.

The corpus at hands consisted of 3,767 German customer reviews, which have been extracted from 38 different social media platforms. Three subjects were asked to annotate the sentiment of each given sentence (positive, neutral, negative) while labelling the aspect (the item that is evaluated, e.g., “room”).

Developing a supervised machine-learning algorithm, we focus on the use of linguistic salience. Our approach comprises a detailed analysis of the language used within the reviews taking into consideration the differences with respect to the distribution of part-of-speech (POS) tags between product reviews and service reviews, as well as differences occurring with respect to the sentiment.

Concerning the annotated aspects, nouns and proper nouns constitute 82% of all POS tags:

- “Das Zimmer jedoch war sehr enttäuschend.“ (*Engl.* “However, the room was very disappointing.”)
- “Nutella ist super lecker.“ (*Engl.* “Nutella is super yummy.”)

The aspect category occurring second most was pronouns (5.5%):

- “Alles perfekt.“ (*Engl.* “Everything perfect.”)
- “Er war noch pampig zu mir.“ (*Engl.* “He even was sloppy with me.”)

This raises the question of when and why customers choose pronouns instead of nouns or proper nouns to function as aspects within an evaluative sentence.

Within our paper, we will examine the role of the following two influencing factors on aspect choice:

1) The sentiment of the evaluative sentence (positive, neutral, negative)

2) The category of the evaluated aspect (product, service)

In the following, Section 2 deals with related literature covering psychological and linguistic research relating to the use of pronouns including sentiment analysis. In Section 3, the procedure of the annotation study is presented, leading to the presentation and discussion of results in Section 4. Finally, Section 5 finishes this contribution with a brief conclusion and information about future research questions.

II. LINGUISTIC AND PSYCHOLOGICAL RESEARCH

As opposed to psychologically derived linguistic dimensions such as emotion words, pronouns constitute a closed class of standard grammatical units [6] [7]. The use of pronouns in verbal behavior has been studied from various linguistic and psychological viewpoints.

Substituting an antecedent noun or noun phrase, pronouns establish coherence in a text [8]. Crawley et al. [9] state that there is a preference for pronominal coreference if the antecedent is mentioned prominently (subjecthood, first position, clefted phrase). Bosch and Umbach [10] found out that for non-subject antecedents, the preference for demonstrative pronouns is stronger than for personal pronouns with subject antecedents.

With regard to personality measures, findings suggest that the use of first person singular pronouns (e.g., I, me, my) correlates with the degree of self-involvement and self-awareness of the speaker [11] [12]. A person who is focusing attention on himself is more likely to use first person singular pronouns [13]. For example, depression was found to be related to an increased use of first person singular pronouns in patients, reflecting a weakness in connecting to others [14]. However, the more the health of the patients improved, the more the individuals shifted in their use from first person to second and third person pronouns [15].

Following Sillars et al. [16], married couples who live in a satisfying and long lasting relationship are more likely to use first person plural pronouns (e.g., we, us, our) to demonstrate their sense of belonging.

Against the background of lying and truth-telling, studies found out that truth-tellers use a higher rate of first person singular pronouns than liars [17]. Liars use less first person singular pronouns to detach themselves from the lie being told [18]. Knowledge about the increased use of first person singular pronouns in truth-telling has led to the fact that liars and fake reviewers nowadays tend to overuse first person singular pronouns to persuade others [18][19].

Relating to research dealing with social interaction and politeness, Brown and Levinson state that in order to keep one's face while conveying a negative message, people impersonalize their opposite by avoiding the pronouns "I" and "you" [20]. In order to motivate others, people formulate statements, which are characterized by a low degree of first person singular pronouns emphasizing the spirit of collaboration needed [21].

Although pronouns carry no sentiment as such in [8], they function as linguistic markers for sentiment analysis.

Ofek et al. [22] used pronouns to predict whether a statement contains objective or subjective content stating that subjective statements contain more personal pronouns than objective statements. Similar Moen et al. [23] found out that when angry, people use more second and third person pronouns to focus on others than themselves. However, studies dealing with the use of pronouns against the background of sentiment analysis do not consider the role attributed to the pronoun within an evaluative sentence.

III. ANNOTATION STUDY

As described in our previous work [24], 38 different social media platforms have been chosen for data extraction. The open-source Java library jsoup [25] was implemented to extract 3,767 German customer reviews relating to products and services. The reviews were parsed into single sentences and tokenized using the Stuttgart-Tübingen TagSet (STTS) [26] and Stanford Parser [27].

Three German native speakers familiar with the process of annotation have been asked to annotate the customer reviews. Subjects were asked to annotate the sentiment of each given sentence (positive, neutral, negative) while labelling the aspect (the item that is evaluated, e.g., "room"). In case of ambiguous or sub-clause sentences, subjects were allowed to mark more than one aspect and sentiment value. For instance, within the sentence "Die Lage ist gut, aber der Raum sehr klein." (*Engl.* "The location is good, but the room is very small."), subjects were able to label two aspects and two sentiment values. The first part of the sentence was labelled positive ("good") relating to the aspect "location" and the second part dealing with the aspect "room" was labelled negative ("small"). This ensures that all information provided in a sentence are gathered.

Prior to annotation, the annotation process was explained to the subjects with three exemplary sentences. The sentences varied with respect to the level of complexity. This guarantees consistency amongst annotators, i.e., it should ameliorate the interrater reliability.

IV. RESULTS AND DISCUSSION

Starting with analyzing the interrater reliability, Fleiss' Kappa values for aspect choice and sentiment values range between 0.5 and 0.8. Thus, they are located in moderate level of agreement [28].

Investigating the influence of sentiment on aspect choice, Figure 1 depicts the distribution of pronouns with respect to the sentiment. As the class for neutral sentiment value is with approximately 5% very low, we will not consider this class in our further analyses. Moreover, previous studies revealed that the class of neutral sentiment has to be examined separately [24].

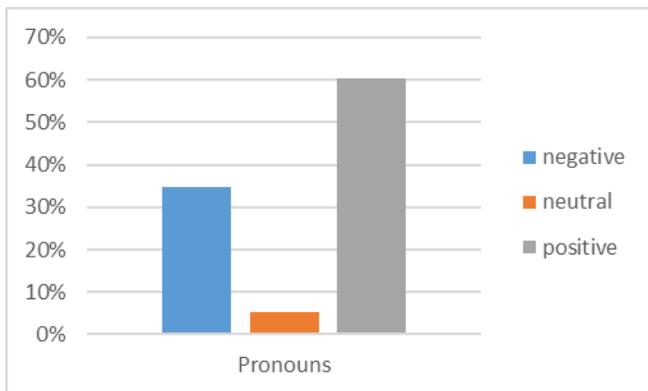


Figure 1. Distribution of pronouns as aspects with respect to the sentiment.

Figure 1 illustrates the distribution of pronouns denoting the aspect of an evaluative sentence against the sentiment of the sentence. Of all pronoun-aspect occurrences, 60% occur in positive sentences. The use of pronouns as aspects tends to be most prominent in evaluative sentences containing a positive sentiment. One first explanation could be that customers use more detailed forms of language, i.e., actual product parts or certain aspects of a service, if they formulate critique or evaluate the product with a negative outcome. In contrast to that, while formulating a positive review, customers tend to generalize.

Figure 2 depicts the distribution of pronouns denoting the aspect of an evaluative sentence against the sentiment of the sentence and the product category. It is striking that within the category service reviews, the number of pronouns-aspect occurrences is equal with respect to the sentiment of the sentence. When evaluating services, the sentiment of the sentence has no influence on the selection of the POS to represent the aspect. However, relating to reviews referring to products there exists a difference in the number of pronouns used as aspects between positive and negative sentiments. If a product is evaluated with a positive sentiment, the aspect is more likely to be realized by a pronoun.

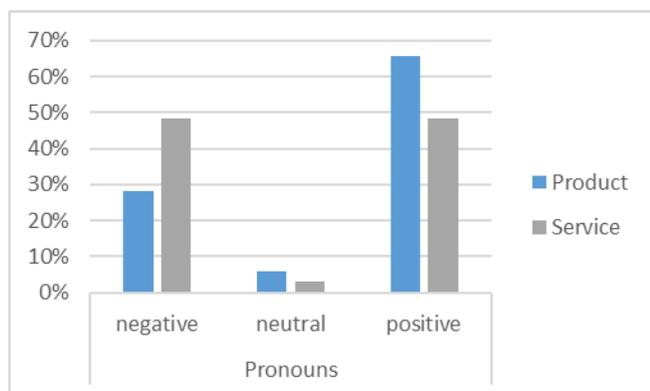


Figure 2. Distribution of pronouns as aspects with respect to sentiment and product category.

Taking a closer look at the actual pronoun tokens, which function as aspects within evaluative sentences, Figure 3

illustrates the words, which have been chosen to represent the aspect of an evaluative sentence. In most cases, 3rd person singular pronouns (“er, sie, es”, “der, die, das”) were attributed the role of the aspect.

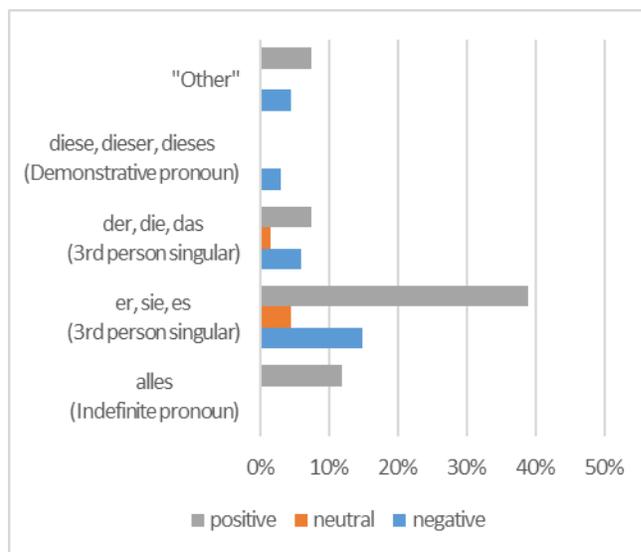


Figure 3. Distribution of used pronouns in product reviews with respect to the sentiment.

However, the indefinite pronoun “alles” (*Engl.*: “everything”) is exclusively used in positive evaluative sentences. Our assumption is that when evaluating products with a positive sentiment, customers differentiate less and summarize their positive impression and experience using indefinite pronouns. On the other hand, when expressing anger within a negative evaluative sentence, customers reflect on single features and details describing precisely what has led to the negative impression and experience. Following this idea we cross-read the sentences labelled as positive and which contain the indefinite pronoun “alles” as aspect. As a result, more than 75% of the sentences contained the statement “everything was good” or words to that effect. Thus, one could argue that these pronouns functioning as aspects in product statements are representative for the general customer satisfaction regarding the given product.

V. CONCLUSION AND FUTURE WORK

The results obtained in our study strengthen the assumption that there exists a difference in the use of pronouns as aspects for negative and for positive evaluative sentences within customer reviews relating to material products. However, customer reviews dealing with services do not show these differences.

Pronouns are more often chosen as aspects if the evaluative sentence is positive. Positive reviews have a tendency to be longer in text length and token number. Therefore, one might argue that pronouns functioning as

aspects are mainly used in longer comments to establish coherence within a text. Although, it is very likely that this argument is not true for the indefinite pronoun “alles” (Engl.: “everything”), it might make sense to investigate the number of sentences per comment, too. In addition, while enlarging the database, the use of pronouns within service reviews should be examined as well.

Whereas there is still a lot of potential for further research from the linguistic point of view within customer reviews in social media, our results already reveal the necessity to use POS information in automatically analyzing these data. Thus, within our objective to develop a supervised machine-learning algorithm to extract relevant information, this information will be taken into account.

ACKNOWLEDGMENT

The support of the German National Science Foundation (DFG) through the funding of the research project “Entwicklung eines Sensors für ungerichtete Beschwerden aus Online Foren” (SCHM 1856/69-1) is gratefully acknowledged.

REFERENCES

- [1] R. Schmitt, S. Schmitt, A. Linder, M. Rüßmann, and V. Heinrichs, “Fehlerinformationen nutzen – Produkte nachhaltig absichern (in English: The usage of failure information – safeguarding products sustainably).” In: 18. Business Forum Qualität. Daten für die Qualität von morgen - generieren, interpretieren und nutzen (in English: 18th Business Forum Quality. Data for tomorrow’s quality – generation, interpretation and usage). September 2014. Apprimus Aachen, pp. 2-22, 2014.
- [2] R. G. Cooper and S. J. Edgett, “Ideation for product innovation: What are the best methods.” PDMA visions magazine 1(1), pp. 12-17, 2008.
- [3] F. T. Piller, A. Vossen, and Ch. Ihl, “From Social Media to Social Product Development: The Impact of Social Media on Co-Creation of Innovation.” Die Unternehmung, 65(1), pp.1 – 22, 2012.
- [4] K. Schouten and F. Frasincar, “Survey on aspect-level sentiment analysis.” IEEE Transactions on Knowledge and Data Engineering 28(3), pp. 813-830, 2016.
- [5] K. Boland, A. Wira-Alam, and R. Messerschmidt, “Creating an annotated corpus for sentiment analysis of german product reviews.” GESIS - Leibniz-Institut für Sozialwissenschaften. Mannheim, GESISTechnical Reports 2013/05, May 2013.
- [6] B. Hall Partee, “Opacity, coreference, and pronouns.” Synthese, 21(3), pp. 359-385, 1970.
- [7] J. W. Pennebaker, M. R. Mehl, and K. G. Niederhoffer. “Psychological aspects of natural language use: Our words, our selves.” Annual review of psychology 54(1), pp. 547-577, 2003.
- [8] H. W. Cowles, “The influence of ‘aboutness’ on pronominal coreference.” ZAS Papers in Linguistics, 48, pp. 23–38, 2007.
- [9] R. A. Crawley, R. J. Stevenson, and D. Kleinman, “The use of heuristic strategies in the interpretation of pronouns.” Journal of Psycholinguistic Research 19(4), pp. 245-264, 1990.
- [10] P. Bosch and C. Umbach, “Reference determination for demonstrative pronouns.” ZAS Papers in Linguistics 48, pp. 39 – 51, 2007.
- [11] W. Ickes, S. Reidhead, and M. Patterson, “Machiavellianism and self-monitoring: as different as ‘me’ and ‘you’” Soc. Cogn. 4, pp. 58–74, 1986.
- [12] L. Scherwitz and J. Canick, “Self reference and coronary heart disease risk.” In: Type A Behavior Pattern: Research, Theory, and Intervention, ed. K Houston, CR Snyder, New York: Wiley, pp. 146–167, 1988.
- [13] D. Davis and T. C. Brock , “Use of first person pronouns as a function of increased objective self-awareness and performance feedback.” J. Exp. Soc. Psychol. 11, pp. 389–400, 1975.
- [14] W. Bucci and N. Freedman, “The language of depression.”, Bull. Menninger Clin. 45, pp. 334–358, 1981.
- [15] R. S. Campbell and J. W. Pennebaker, “The secret life of pronouns: Flexibility in writing style and physical health.” Psychological science 14(1), pp. 60-65, 2003.
- [16] A. Sillars, W. Shellen, A. McIntosh, M. Pomegranate, “Relational characteristics of language: elaboration and differentiation in marital conversations.” West. J. Commun. 61, pp.403–422, 1997.
- [17] M. L. Newman, J. W. Pennebaker, D. S. Berry, and J. M. Richards, “Lying words: Predicting deception from linguistic styles.” Personality and social psychology bulletin, 29(5), pp. 665-675, 2003.
- [18] B. Liu, “Sentiment analysis: Mining opinions, sentiments, and emotions”, Cambridge University Press, 2015.
- [19] E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco, “Affective Computing and Sentiment Analysis.” In A Practical Guide to Sentiment Analysis, Springer International Publishing, pp. 1-10, 2017.
- [20] P. Brown, and S. C. Levinson, “Politeness: Some universals in language usage.” Cambridge university press, 1987.
- [21] J. W. Pennebaker & L. A. King, “Linguistic styles: language use as an individual difference.” J. Personal. Soc. Psychol. 77, pp. 1296–1312, 1999.
- [22] N. Ofek, L. Rokach, C. Caragea, and J. Yen, “The Importance of Pronouns to Sentiment Analysis: Online Cancer Survivor Network Case Study.” Proceedings of the 24th International Conference on World Wide Web. ACM, pp. 83 – 84, 2015.
- [23] M.- F. Moens, J. Li, and T.-S. Chua, “Mining user generated content.” CRC Press, 2014.
- [24] A. I. Metzmacher, V. Heinrichs, B. Falk, and R. H. Schmitt, „Use of Negation Markers in German Customer Reviews“, The Eleventh International Conference on Advances in Semantic Processing (SEMAYRO 2017), IARIA, November 2017, pp. 37 - 41.
- [25] J. Hedley “jsoup: Java HTML Parser”, 2009. [Online] Available from: <https://jsoup.org/> [Retrieved: November, 2017].
- [26] Insitut für Maschinelle Sprachverarbeitung, “German Tagsets”, 2013. [Online] <http://www.ims.uni-stuttgart.de/forschung/ressourcen/lexika/GermanTagsets.html> [Retrieved: November 2017].
- [27] A. N. Rafferty and C. D. Manning, “Parsing three German treebanks: Lexicalized and unlexicalized baselines”. In Proceedings of the Workshop on Parsing German. Association for Computational Linguistics, pp. 40-46, 2008.
- [28] M. L. McHugh, “Interrater reliability: the kappa statistic”, Biochemia medica, 22(3), pp. 276-282, 2012.