The Importance of Canonical Product Attributes on User Opinions: an Empirical Investigation

Tiago de Melo
Universidade do Estado do Amazonas
Manaus, Brazil
tmelo@uea.edu.br

Altigran S. da Silva
Universidade Federal do Amazonas
Manaus, Brazil
alti@icomp.ufam.edu.br

Edleno S. de Moura
Universidade Federal do Amazonas
Manaus, Brazil
edleno@icomp.ufam.edu.br

Pável Calado
Universidade de Lisboa
Lisboa, Portugal
pavel.calado@tecnico.ulisboa.pt

ABSTRACT

When making purchasing decisions, customers usually rely on information from two types of sources: product specifications, provided by manufacturers, and reviews, posted by other customers. Both kinds of information are often available on e-commerce websites. While researchers have demonstrated the importance of product specifications and reviews as separate and valuable sources to support purchase decision-making, a largely uninvestigated issue is what is the relationship between these two kinds of information. In this paper we present an empirical study on the use of direct and indirect mentions to canonical product attributes, that is, those defined by manufactures in product specifications, in the reviews written by customers. For this study, we analyzed more than 1,100,000 opinionated sentences available in about 650,000 user reviews from Amazon.com across five product categories. Our results indicate that user opinions are indeed guided by the attributes from product specifications and highlight the influence of canonical attributes on the user reviews.

CCS CONCEPTS

• **Information systems → Sentiment Analysis**; *Web mining*.

KEYWORDS

User Reviews, Product Attributes, Opinion Mining

ACM Reference Format:

Tiago de Melo, Altigran S. da Silva, Edleno S. de Moura, and Pável Calado. 2019. The Importance of Canonical Product Attributes on User Opinions: an Empirical Investigation. In *Companion Proceedings of the 2019 World Wide Web Conference (WWW'19 Companion), May 13–17, 2019, San Francisco, CA, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3308558.3316458

WWW '19 Companion, May 13–17, 2019, San Francisco, CA, USA

ACM ISBN 978-1-4503-6675-5/19/05.
https://doi.org/10.1145/3308558.3316458

1 INTRODUCTION

In typical e-commerce Web sites, descriptions of products in the catalog usually consist of objective (factual) data provided by manufacturers informing customers of product's characteristics, which are represented as a set of previously defined product attributes. For instance, for laptops, the brand, the weight and the processor model are commonly available to help potential customers make their purchase decisions. These attributes are called here *canonical*. However, with the rise of the so-called Web 2.0, there is also a large amount of subjective (opinionated) information available on products and their characteristics. In most cases, this subjective information is provided by opinions issued by other customers in reviews.

The importance of considering subjective information in addition to objective (factual) information has been verified in many e-commerce related applications. Indeed, considering opinions issued by other people before purchasing a product is a common practice, especially since there are plenty of opinions available on the Web. In a recent survey [15], it was found that 82% of the Americans refer to online reviews when they first purchase a product, and 40% always refer to online reviews when purchasing good. Another comprehensive survey of online shoppers from 5 different continents [10] revealed that 45% of consumers consider reviews as the most influential aspect of social media for their online shopping behavior. This survey also shows that checking reviews about products or retailers is the fourth most common activity of in-store shoppers with mobiles/smartphones.

On the other hand, product attributes also comprise valuable sources to support purchase decision-making. According to Park et al. [8], product attributes in websites encourage consumer browsing behavior, which can often lead to impulse buying behavior. Therefore, product attributes are a crucial element that influences customer product choice [2].

Despite the substantive importance of user opinions and product attributes for customer decisions, the relationship between these two kinds of information has been largely overlooked. Motivated by the above observations, in this paper, we empirically study the importance of product attributes on user opinions. Specifically, we analyze the use of direct and indirect mentions to canonical product attributes, that is, those defined by manufactures in product specifications, in the reviews written by customers. Our ultimate goal is to extend previous research studies that assessed the importance of product attributes and user opinions in a separate way. To this end, we executed an extensive experimental evaluation using about 650,000 real reviews, composed of more than 1,100,000 opinionated sentences, and more than 30,000 products in five different categories of electronic products (cameras, cell phones, DVDs, laptops, and routers). The results from this study revealed the importance of canonical product attributes on the user opinions.

The remainder of this paper is organized as follows. Section 2 provides a brief review of related work. In Section 3, we describe hypotheses we have formulated for this study. Then, we describe the experimental dataset we used to verify our hypotheses in Section 4 and the main results and findings that support our hypotheses in Section 5. Finally, we conclude by discussing our results and by proposing directions for future research in Section 6.

2 RELATED WORK

In this section, we describe related work been done on the importance of product attributes [3, 6, 13, 17] and user reviews on purchasing decisions [1, 2, 12, 14, 16].

There are some studies which investigated the importance of product attributes in e-commerce domain. Lee and Nguyen [3] investigated the importance of product attributes in purchasing fashion goods and the influence of these attributes on preferences for external versus local fashion brands. Rahul Rai [13] proposed a method that identifies key product attributes from online customer reviews and ranks each product's attributes with part-of-speech (POS) tagging. Maslowska et al. [6] studied how review characteristics (i.e., number of reviews), product characteristics (i.e., price) and customer behaviors (i.e., reading reviews) interact with each other to influence purchase decisions. Wang et al. [17] conducted a study on online reviews to measure how product attributes impact customer satisfaction.

On the other hand, several works have investigated the importance of customer reviews. Kostyra et al. [2] investigated the impact of online customer reviews on customer's decisions. Qazi et al. [12] investigated why some reviews are more helpful compared to others. Jo and Oh [1] proposed two distinct models to discovering what aspect expressions are evaluated in user reviews and how sentiments for different aspects are expressed. Sun et al. [16] proposed a method of using external user generated data to evaluate the relative importance of an entity's attribute. Singh et al. [14] developed a method based on machine learning that can predict the helpfulness of the consumer reviews using several textual features such as polarity, subjectivity, entropy, and reading ease.

While prior studies have contributed to understanding on the importance of product attributes and user reviews as separate and valuable sources to support purchase decision-making, our work investigated the relationship between two kinds of information.

3 RESEARCH HYPOTHESES

In this study, we aim at investigating the importance of canonical product attributes on user opinions. We choose online user reviews on electronic products available in e-commerce websites as a context for our research due to the popularity and importance of this domain.

A review is a text posted by a user on an e-commerce website, usually reporting their experience with a specific product, which we call the *target product* of the review. Each review is composed of a set of *sentences*. Sentences that express factual information are called *objective* sentences, while sentences that express personal feelings or beliefs are called *subjective* or *opinionated* sentences [5]. We are interested in the latter because they represent the user's opinions on a product.

An opinionated sentence can be further classified as *comparative* or *direct*. A comparative sentence expresses an opinion on similarities or differences between two or more products. The sentence *"the camera of the iPhone is much better than Galaxy"* is an example of a comparative sentence. A direct opinionated sentence (DOS) expresses an opinion directly on a characteristic or part of the product, or on the product as a whole. The sentence *"the camera of the iPhone is fantastic"* is an example of a direct opinion. As our goal is to investigate the importance of canonical product attributes regarding the *target product*, we only consider direct opinionated sentences, since opinions in comparative sentences are relative.

For online user reviews posted on e-commerce websites, we expect to obtain a huge volume of DOS. We thus formulate the following hypotheses:

$\mathcal{H}1$**.** E-commerce websites are a valuable source of opinions on *target products*.

When analyzing real user reviews, we noticed that reviews also include opinions that do not refer to a canonical product attribute, but to the product as a whole. Furthermore, opinions may also refer to attributes that are not represented in the product catalog. Thus, we consider that opinions may have three distinct targets: a) `Attribute`, when opinion is on a canonical product attribute; b) `General`, when opinion is on the product as whole; and c) `Other`, when opinion is on a characteristic of the target product that is not represented as a canonical attribute. In spite of that, we do expect that canonical attributes have a persuasive effect on online user reviews. Hence, we hypothesize:

$\mathcal{H}2$**.** Most of the user opinions posted in e-commerce web sites is on canonical product attributes.

$\mathcal{H}3$**.** According to the user opinions, there are certain canonical attributes that are more relevant than the other attributes.

$\mathcal{H}4$**.** Customers often make either direct and indirect mentions to canonical product attributes.

To verify our hypotheses, we built an experimental dataset that will be explained in the next section.

4 EXPERIMENTAL DATASET

We built our experimental dataset from a large collection of about 142 million reviews previously crawled from the Amazon.com web site [7][1], and selected all reviews from each of five categories: cameras (CAM), cell phones (CEL), dvd players (DVD), laptops (LAP), and internet routers (ROT). Each review in this collection identifies the product to which it refers. Table 1 presents, for each category, the number of reviews and sentences in this collection, along with the number of products referred in the reviews.

[1]Available at http://jmcauley.ucsd.edu/data/amazon

Table 1: Summary of Amazon.com review collection.

Category	No. products	No. reviews	No. sentences
CAM	8,893	203,836	1,012,077
CEL	7,693	182,491	707,407
DVD	2,503	61,836	243,939
LAP	9,491	115,138	580,955
ROT	1,592	84,059	329,305
Total	**30,172**	**647,360**	**2,864,683**

To complete our experimental dataset, we crawled from Amazon.com the set of canonical attributes used in products of each of the five categories. These sets are shown in Table 2.

Table 2: Canonical product attributes.

Category	Product Attributes
CAM	*Dimension*, *Exposure Control*, *Imaging Memory*, *Performance*, *Power*, *Price*, *Zoom*
CEL	*Battery*, *Camera*, *Dimension*, *Display Memory*, *Price*, *Processor*, *Software*
DVD	*Accessory*, *Audio*, *Dimension*, *Price Sound*, *Video*
LAP	*Battery*, *Connectivity*, *Dimension*, *Graphic Memory*, *Price*, *Processor*, *Screen*, *Software*
ROT	*Accessory*, *Coverage Area*, *Dimension Ports*, *Price*, *Security*, *Software*, *Speed*

Recall that our study targets only direct opinionated sentences (DOSs). Thus, we proposed a method named *filterDOS* to select DOSs from reviews.

filterDOS is carried out in three steps and it allows to identify three types of sentences: *factual*, *comparative*, and *DOS*. In the first step, the reviews are broken down into sentences. In the second step, the method identifies subjective sentences obtained in the first step. Objective (factual) sentences are discarded. As an example, the sentence "*I had to go back to store and they did not have the device stocked*" written by a user is not opinionated and therefore it should be discarded. For this task, we implemented the unsupervised method proposed by Qadir [11], where typed dependency relations, such as open clausal complements or adjectival complements, are used for identifying subjective sentences. The subjective sentences can be further classified into *comparative* or *direct* sentence. A comparative sentence expresses a relation of similarities or differences between two or more products. As an example, the sentence "*the camera of the iPhone is much better than Galaxy*" is a comparative sentence. A direct opinionated sentence (DOS) expresses an opinion directly about a characteristic or part of the product, or the product as whole. The sentence "*the camera of the iPhone is fantastic* is an example of DOS. Finally, in the last step, *filterDOS* eliminates the comparative sentences, based on the unsupervised method proposed by Liu [4].

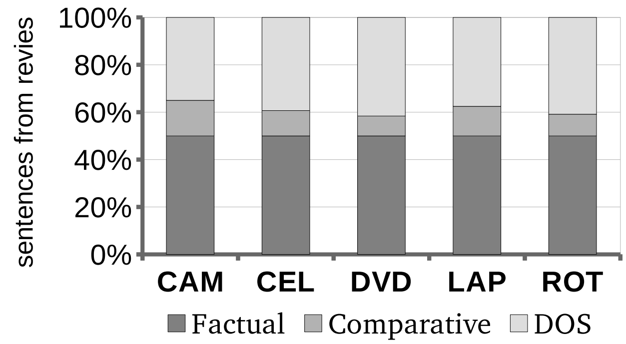


Figure 1: Percentage (%) of sentences of each type in user reviews.

Figure 1 presents the results obtained using *filterDOS* on the Amazon.com review collection. We can observe that, in average, 51.31% of the sentences are factual, 40.05% are DOSs, and only 8.64% are comparative. It is noticeable that there are very few of comparative sentences in product reviews. This is due to the fact that e-commerce site users focus on writing only about the product of interest, unlike what occurs, for example, in forums where users usually write comments comparing products.

The initial review collection described in Table 2 has more than 2,800,000 sentences. Therefore, even after filtering factual and comparative sentences, the number of sentences to be analyzed is still huge, with more than a one million sentences.

A central concept in opinion mining that we leverage in our study is that of *aspect expressions*. We consider that an opinion is represented by an aspect expression [5], where an aspect is any reference made in an opinion to a particular part or characteristic of the product, or even to the product as a whole. Since it would be unfeasible to manually annotate all aspects found this huge volume of sentences, we created an experimental dataset focused on the 100 most frequent aspect expressions found in the DOS of each product category. We argue that in a practical setting, handling a few top frequent aspect expressions is more valuable than considering every single aspect expression from a potentially huge list.

To select the 100 most frequent aspect expressions, we first run the aspect extraction method proposed by Poria et al. [9]. This was implemented and all possible aspect expressions identified by this method were extracted. We then ranked these expressions according to their frequency. To assure that we only use true aspect expressions, we manually inspected the extracted expressions using the ranking order, and removed those that we did not consider as aspect expressions. In the end, only the 100 most frequent true aspect expressions were kept for each product category. We named this set of 100 aspect expressions as *top 100 aspects*.

Finally, for each product category, we manually examined each of the *top 100 aspects* selected previously, and annotate each one with the canonical product attributes that is most related to it. We also annotate accordingly cases where the opinion is on the product as whole (`General`) and when opinion is on a characteristic of the target product that is not represented as canonical attribute (`Other`). This annotated dataset was then used to verify our formulated hypotheses ($\mathcal{H}1$, $\mathcal{H}2$, $\mathcal{H}3$, and $\mathcal{H}4$). Table 3 presents a summary of the experimental dataset we generated, which shows, for each category, the total number of DOSs (No. DOSs) and the total number of DOSs that include at least one of the top 100 aspects (No. DOSs *top100*). As it can be observed, these DOSs represent more the 50% of all DOSs.

Table 3: Summary of our experimental dataset.

Category	No. DOSs	No. DOSs *top100*
CAM	476,605	249,714
CEL	277,712	138,939
DVD	89,525	48,608
LAP	189,782	126,865
ROT	123,336	73,027
Total	**1,156,960**	**637,153**

5 RESULTS

In this section, we present the main results and findings that support our hypotheses formulated on the importance of canonical product attributes on user opinions.

5.1 Distribution of Sentences among Kinds of Targets

To test the hypothesis $\mathcal{H}1$, we created the method *filterDOS* to select direct opinionated sentences from reviews. According to our experimental dataset presented in Table 3, more than 1,100,00 sentences are DOSs. The fact that a large fraction of the sentences is DOS supports our hypothesis $\mathcal{H}1$ that e-commerce websites, such as Amazon.com, are indeed useful as a valuable source of opinions on target products. Furthermore, from the experimental dataset, we can observe that more than 55% DOSs contain at least one of the *top 100 aspects*. This finding corroborates our assumption that handling a few top frequent aspect expression is more valuable than showing every single aspect expression from a potentially huge list.

Figure 2 summarizes the distribution of sentences among the three kinds of targets: `Attributes`, `General`, and `Other`. The number of sentences containing at least one of the *top 100 aspects* that form the opinions are annotated with a kind of target. As explained in Section 3, a single sentence may contain more than one opinion, and each opinion can refer to a different kind of target. Thus, the sum of percentage of all kinds of targets may be greater than 100%.

Again we observe that most of the DOSs include aspect expressions that refer to canonical product attributes, identified as `Attribute`. For example, in CAM and LAP categories they account for more than half of the sentences. Also, a large share of the sentences contain opinions referring to the target `General`.

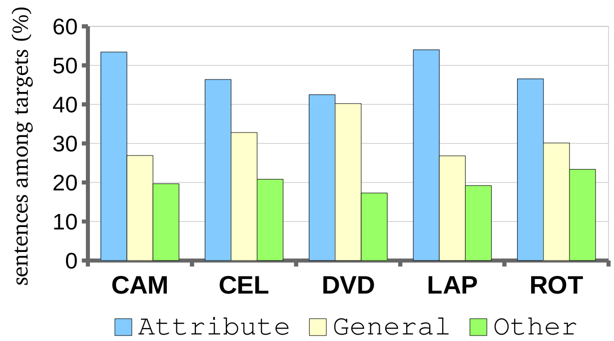


Figure 2: Distribution of sentences among targets.

5.2 Distribution of Aspect Expressions

Figure 3 shows the distribution of the *top 100 aspects* among the three kinds of targets. Notice that for all categories the fraction of aspect expressions that represent the target `Attribute` is higher than 50%. This supports our hypothesis $\mathcal{H}2$, that most of the user opinions posted in e-commerce web sites is about the canonical product attributes. In addition, it can be observed that a larger share of the aspect expressions refer to the target `Other`. For example, in CAM, CEL and DVD, almost one quarter of the aspect expressions are in opinions which were annotated as the target `Other`.

An intriguing problem we left for future work is to further analyze cases such as these to look for specific latent characteristics that, although are not represented by some canonical attribute, are of interest for users. For example, "keyboard" is the second most frequent aspect expression in LAP, but typically, there is no attribute referring to it in the canonical product attributes.

In sum, Figure 3 suggests that users comment more frequently on the specific characteristics of the products than on the product as a whole. This shows the relevance of properly addressing references to attributes in user reviews.

5.3 Distribution of Sentences among Canonical Product Attributes

Figure 4 shows the distribution of sentences among the canonical product attributes for each category. In these graphs, each vertex in the polygon represents a canonical product attribute defined by manufacturers in product specifications. The graph shows the percentage of sentences that contain an aspect expression that corresponds to a given attribute. In each graph, the canonical attributes are placed in clockwise order, from the most to the least frequently referred. For example, more than 40% of the sentences that include at least one of the *top 100 aspects* in the DVD category refer to the attribute *Accessory*.

There are some attributes that are much more frequently referred to in reviews than other from the same category. For example, in the CEL category, users comment 2 times more on `Battery` than on the `Price` of cell phones. These experimental results support our hypothesis $\mathcal{H}3$ that there are certain canonical attributes that are more relevant than the other. Interestingly, in the five categories in this experiment, the price is not the most commented attribute.

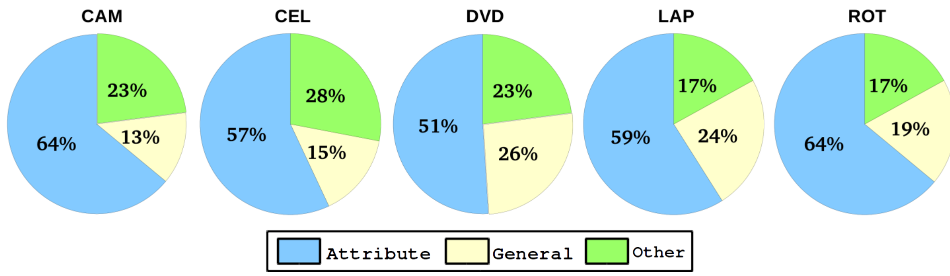
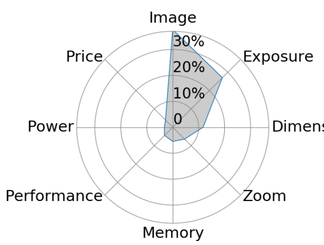
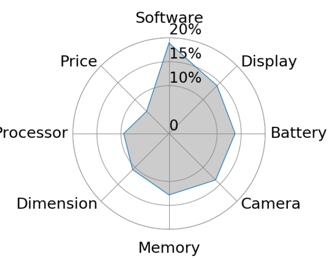


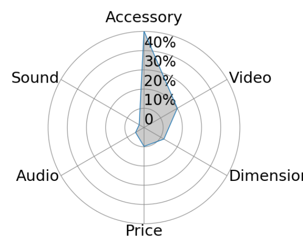
Figure 3: Distribution of the top 100 aspects among the three kinds of targets.



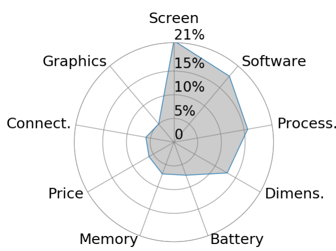
(a) CAM



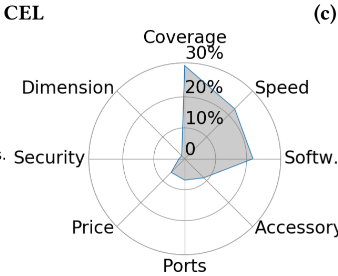
(b) CEL



(c) DVD



(d) LAP



(e) ROT

Figure 4: Distribution of sentences among the attributes they refer to. Labels are canonical product attributes.

5.4 Diversity of Aspect Expressions over Attributes

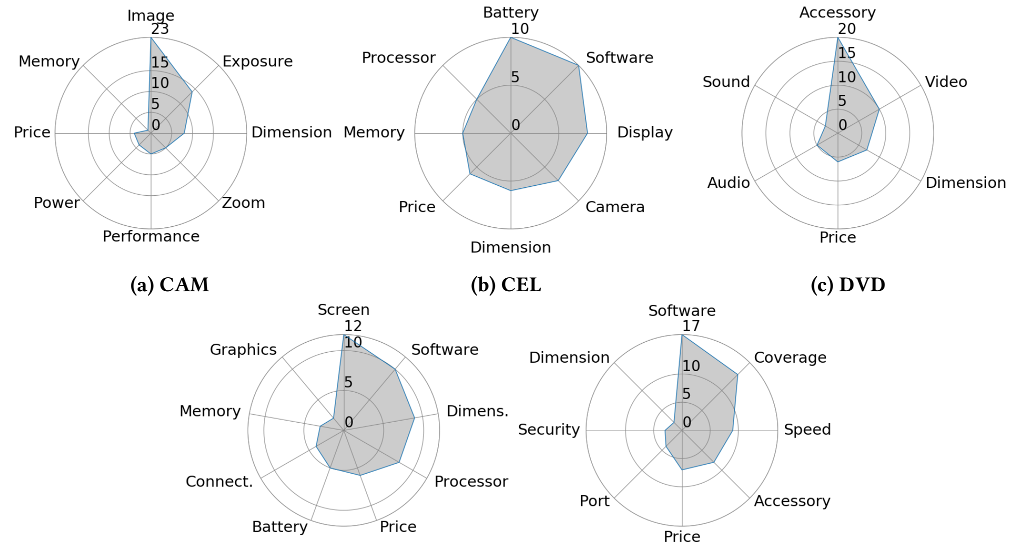
Figure 5 shows the distribution of aspect expressions extracted from user reviews over canonical product attributes in each category. In these graphs, we show the quantity of unique aspect expressions that refer to the same attribute. For example, in the LAP category, we found ten different aspect expressions that refer to the attribute *Software*. Analyzing the sentences, we found that users do indeed employ several different terms such as “apps”, “system”, “vista”, and “program” to refer the attribute *Software* in the LAP category. This experimental evaluation supports our hypothesis $\mathcal{H}4$ that customers often make either direct and indirect mentions to canonical product attributes.

To give an idea of the *top 100 aspects* mentioned, Table 4 illustrates the ten most frequent aspect expressions extracted in the reviews of each category, along with the attribute name, when they refer to `Attribute` target, or the target name (`General` or `Other`). From these results is it apparent that the ten most frequent aspect expressions extracted are quite representative of each product category and, more importantly, the results show which are the most commented aspects related to canonical attributes. Notice that, most aspect expression do not match exactly with the name of the product attribute.

6 CONCLUSIONS

The main goal of this research is to investigate the importance of canonical product attributes on user opinions. Based on a empirical evaluation carried out over a representative collection of real user reviews, we were able to verify hypotheses we have formulated on this issue. Our results were drawn from a large experimental dataset we built with more than 1 million direct opinionated sentences (DOSs) in five product categories.

In our study, we verify that a large fraction of sentences in reviews is composed of direct opinionated sentences and this validate our hypothesis that e-commerce websites are a valuable source of opinions on target products ($\mathcal{H}1$). We use this large number of sentences and, by means of a well defined protocol, we could verify that the most of user opinions posted in e-commerce web sites is on one of the canonical product attributes ($\mathcal{H}2$). Furthermore, we could verify that there are certain canonical attributes that are



(a) CAM **(b) CEL** **(c) DVD**

(d) LAP **(e) ROT**

Figure 5: Distribution of different aspect expressions according the canonical product attributes presented in Table 2.

Table 4: The 10 most frequent aspect expressions in reviews of each category from the Amazon dataset.

CAM	CEL	DVD	LAP	ROT
camera (General)	phone (General)	unit (General)	laptop (General)	router (General)
quality (General)	easy (Other)	dvd (General)	keyboard (Other)	easy (Other)
picture *(Imaging)*	quality (General)	easy (Other)	computer (General)	instructions (Other)
lens *(Exposure Control)*	feature (Other)	quality (General)	software *(Software)*	software *(Software)*
shots *(Exposure Control)*	card *(Memory)*	player (General)	fast *(Processor)*	unit (General)
features (Other)	size *(Dimension)*	picture quality *(Video)*	size *(Dimension)*	speed *(Speed)*
size *(Dimension)*	software *(Software)*	instructions (Other)	card *(Memory)*	device (General)
card *(Memory)*	camera *(Camera)*	features (Other)	screen *(Screen)*	internet *(Coverage Area)*
settings (General)	screen *(Display)*	picture *(Video)*	easy (Other)	network *(Coverage Area)*
pics *(Imaging)*	keyboard (Other)	product (General)	graphics *(Graphics)*	settings (Other)

more relevant for users than the other attributes ($\mathcal{H}3$). Finally, we could conclude that customers often make either direct and indirect mentions to canonical product attributes using several distinct expressions ($\mathcal{H}4$).

This study contributes to understanding the importance of canonical product attributes on user opinions and our results indicate that user opinions are indeed guided by the attributes from product specifications and highlight the influence of canonical attributes on the user reviews.

Some limitations are associated with this study, which, however, can provide directions for future research. First, we considered only aspect expressions to represent user opinions. Future research could extend the current study for examining other components of opinions, such as star ratings of reviews, opinion polarity, and opinion posting time. Second, our analysis is restricted to products, which in turn, have well established set of canonical attributes provided by manufacturers. However, domains such as restaurants or hotels do not have clear canonical attributes. Therefore, a future research could extend the current study to these domains.

REFERENCES

[1] Yohan Jo and Alice H Oh. 2011. Aspect and sentiment unification model for online review analysis. In *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 815–824.

[2] Daniel S Kostyra, Jochen Reiner, Martin Natter, and Daniel Klapper. 2016. Decomposing the effects of online customer reviews on brand, price, and product attributes. *International Journal of Research in Marketing* 33, 1 (2016), 11–26.

[3] Jaeha Lee and Minhthu Jill Nguyen. 2017. Product attributes and preference for foreign brands among Vietnamese consumers. *Journal of Retailing and Consumer Services* 35 (2017), 76–83.

[4] Bing Liu. 2010. Sentiment Analysis and Subjectivity. *Handbook of Natural Language Processing* 2 (2010), 627–666.

[5] Bing Liu. 2015. *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press.

[6] Ewa Maslowska, Edward C Malthouse, and Vijay Viswanathan. 2017. Do customer reviews drive purchase decisions? The moderating roles of review exposure and price. *Decision Support Systems* 98 (2017), 1–9.

[7] Julian McAuley, Rahul Pandey, and Jure Leskovec. 2015. Inferring networks of substitutable and complementary products. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 785–794.

[8] Eun Joo Park, Eun Young Kim, Venessa Martin Funches, and William Foxx. 2012. Apparel product attributes, web browsing, and e-impulse buying on shopping websites. *Journal of Business Research* 65, 11 (2012), 1583–1589.

[9] Soujanya Poria, Erik Cambria, Lun-Wei Ku, Chen Gui, and Alexander Gelbukh. 2014. A Rule-Based Approach to Aspect Extraction from Product Reviews. In *Proceedings of the Second Workshop on Natural Language Processing for Social Media (SocialNLP)*. 28.

[10] PwC. 2016. They say they want a revolution – Total Retail 2016. http://pwc.to/2kRAxYL. Accessed 08 September 2018.

[11] Ashequl Qadir. 2009. Detecting opinion sentences specific to product features in customer reviews using typed dependency relations. In *Proceedings of the Workshop on Events in Emerging Text Types*. Association for Computational Linguistics, 38–43.

[12] Aika Qazi, Karim Bux Shah Syed, Ram Gopal Raj, Erik Cambria, Muhammad Tahir, and Daniyal Alghazzawi. 2016. A concept-level approach to the analysis of online review helpfulness. *Computers in Human Behavior* 58 (2016), 75–81.

[13] Rahul Rai. 2012. Identifying key product attributes and their importance levels from online customer reviews. In *ASME 2012 international design engineering technical conferences and computers and information in engineering conference*. American Society of Mechanical Engineers, 533–540.

[14] Jyoti Prakash Singh, Seda Irani, Nripendra P Rana, Yogesh K Dwivedi, Sunil Saumya, and Pradeep Kumar Roy. 2017. Predicting the âĂIJhelpfulnessâĂİ of online consumer reviews. *Journal of Business Research* 70 (2017), 346–355.

[15] Aaron Smith and Monica Anderson. 2016. Online shopping and e-commerce. https://www.pwc.es/es/publicaciones/retail-y-consumo/assets/total-retail-2016.pdf. Accessed 10 September 2018.

[16] Shengjie Sun, Dong Yang, Hongchun Zhang, Yanxu Chen, Chao Wei, Xiaonan Meng, and Yi Hu. 2018. Important Attribute Identification in Knowledge Graph. *arXiv preprint arXiv:1810.05320* (2018).

[17] Yuren Wang, Xin Lu, and Yuejin Tan. 2018. Impact of product attributes on customer satisfaction: An analysis of online reviews for washing machines. *Electronic Commerce Research and Applications* 29 (2018), 1–11.