

# Tell Me What to Do! Computer-Synthesized Speech as a Persuasive Technology in the Context of Online Reviews

*Emergent Research Forum papers*

**Marcin Lukasz Bartosiak**  
University of Pavia  
ml.bartosiak@universitadipavia.it

**Gabriele Piccoli**  
Louisiana State University  
University of Pavia  
gpiccoli@cct.lsu.edu

## Abstract

Our work focuses on synthesized speech as a persuasive communication tool in the e-commerce and online recommendations context. Persuasion is one of the distinctive elements of human communication and it is present also when computers and the Internet mediate individuals' relations. To understand the mechanisms of online persuasion, we look at a specific type of content: online reviews. The current standard for delivering them to customers is the written word, sometimes accompanied by images. However, their persuasive potential may increase when an online review system uses synthesized speech to convey the content. We first examine how the quality of the message arguments affects the persuasiveness of online recommendations. Second, we investigate whether synthesized speech as a method of online reviews delivery moderates the relationship between online content and individual's behavior.

## Keywords

Persuasion, online reviews, computer-synthesized speech, experimentation.

## Introduction

Our work focuses on computer-synthesized speech as a persuasive communication tool in the e-commerce and online recommendations context. While limited research to date has had this focus, soon we will be interacting with computers via speech regularly (Bajarin 2011). Popular and emerging devices, such as smartphones, smart-watches, heads-up devices (e.g., Microsoft HoloLens) and wearables in general leverage, and increasingly require, vocal interaction.

Persuasion is one of the distinctive elements of human communication. Information Systems research has recently recognized the importance of computer mediated persuasion (Benbasat 2010, O'Keefe 2004, Tarning and Oinas-Kukkonen 2009). More generally, persuasive devices like Fitbit, Misfit, BitBite or Pavlok are quickly being introduced, and information systems scholars will be increasingly called upon to help explain the role of information technology in human persuasion. In this paper we focus on online content persuasiveness and define it as the extent to which information delivered through an IT channel influences a receiver to perform an action expected by the author of the content.

To understand the mechanisms of online persuasion, we look at a specific type of content: online reviews. Reviews are peer-generated evaluations posted on company or third party websites (Mudambi and Schuff 2010). According to the Global Trust Survey the majority of online consumers nowadays declare they trust online reviews as much as personal recommendations (*Global Trust in Advertising and Brand Messages* 2013). In case of some industries, like hospitality, online reviews are the most influential factor in decision-making, contributing significantly to individual's final hotel choice (McGuire 2013, 2014). Thus,

they hold potential ‘to form or modify people’s attitudes about a product’ (Hong and Park 2012, p. 906), making it a suitable context for our research.

The current standard for delivering online reviews to customers is the written word, sometimes accompanied by images. However, their persuasive potential may increase when an online reviews provider uses computer-synthesized speech to communicate the content. Although computerized speech synthesis traces its roots to the early 1900s (Mattingly 1970), only in recent years computer-synthesized speech solutions have become standard features in consumer technology products. Apple, Google, Microsoft and Amazon all have made significant strides in voice recognition and human computer interaction via speech synthesis. Researchers proved that this kind of communication indeed changes people’s attitude towards computers (Nass and Gong 2000; Nass and Lee 2001). Yet, knowledge about the topic is still limited and little is known about the potential effect of computer-synthesized speech in the context of e-commerce transactions. Some notable examples of research in this area focus on computer-synthesized speech in customer service (Qiu and Benbasat 2005) or retailer’s avatars (Wang et al. 2007), but not on online reviews.

We will first examine how different characteristics of the message affect the persuasiveness of online recommendations. Second, we will investigate whether computer-synthesized speech as a method of online reviews delivery affects users’ behavior.

## Literature Review

### *Online Reviews as Persuasive Text*

Positive online rating and reviews can ‘modify people’s attitudes about a product to which the online reviews pertain’ (Hong and Park 2012, p. 906) and strongly influence other consumers’ buying decisions (Goldenberg et al. 2001). Previous research focuses on persuasion by measuring the quantitative elements of online reviews (e.g. Duan et al. 2008; Park et al. 2007) or secondary text characteristics (e.g. Cheung et al. 2012; Yin et al. 2014), but not the content of the message itself. Yet, text feedback influences seller’s credibility over and above numerical ratings (Pavlou and Dimoka 2006).

The importance of text is consistent with persuasion literature, in which argument quality is theorized as a central driver of influence (e.g. Bhattacharjee and Sanford 2006; Petty and Cacioppo 1986). Toulmin (2003) presented a model to describe high quality, reliable arguments in which the main statement – *claim*, is backed by other facts. Such a construction makes the argument more reasonable and makes people more willing to accept it. Although, Toulmin’s model includes more elements, in this paper we focus only on the most common ones. Previous e-commerce research shows that using descriptions with *claim* – the main idea of the argument, supported by *data* and *backing* justifying the claim, increases consumers’ trust in e-commerce websites (Kim and Benbasat 2006), especially when items under consideration are expensive (Kim and Benbasat 2009). The same effect applies to online reviews. When the quality of online review content is high and there is lack of non-verbal cues people form trust beliefs on the basis of text characteristics (Racherla et al. 2012).

### *Computer-Synthesized Speech as a Persuasive Technology*

Previous research has investigated the effects of computer-synthesized speech on human-computer interactions. People behave differently and hold different attitudes towards technology when computers deliver a message via speech versus text. Computer-synthesized speech has been shown to increase credibility, express personality (Nass and Lee 2001) and persuade users (Joo and Lee 2014).

The psychology of speech processing is the theoretical underpinning of these findings (Nass and Gong 2000; Nass and Lee 2001). Recognition of speech, even computer-synthesized speech, is automatic and humans process it unconsciously. Presence of social characteristics in a speech makes people behave different, based on the unconscious belief that only other human-beings may produce speech-like sounds. Thus, the human brain extends the understanding of speech also to computer-synthesized speech and starts looking for social cues of communication as if it was interacting with another person. When people hear computers “speaking,” they ‘make attributions about voice systems using the same rules and heuristics they would normally apply to other humans’ (Nass and Gong 2000, p. 38). The human brain reads these implicit social cues, and forms the belief that the computer is another member of the society.

In response individuals react as if the machines were another social actors (Nass et al. 1994) and start following social norms typical of communication processes between people (Cialdini and Trost 1998). Thus, when a computer produces speech, people unconsciously treat it as more akin to them, not like a machine. This creates the effect of similarity. Since the similarity between individuals increases the persuasiveness of a communication source (Burger et al. 2004), when people start perceiving the computer as a social actor who is similar to them, they are more prone to being persuaded by its messages. Some researchers further claim, that social presence facilitates persuasion and e-commerce sales (Fogg and Tseng 1999; Lee and Nass 2003)

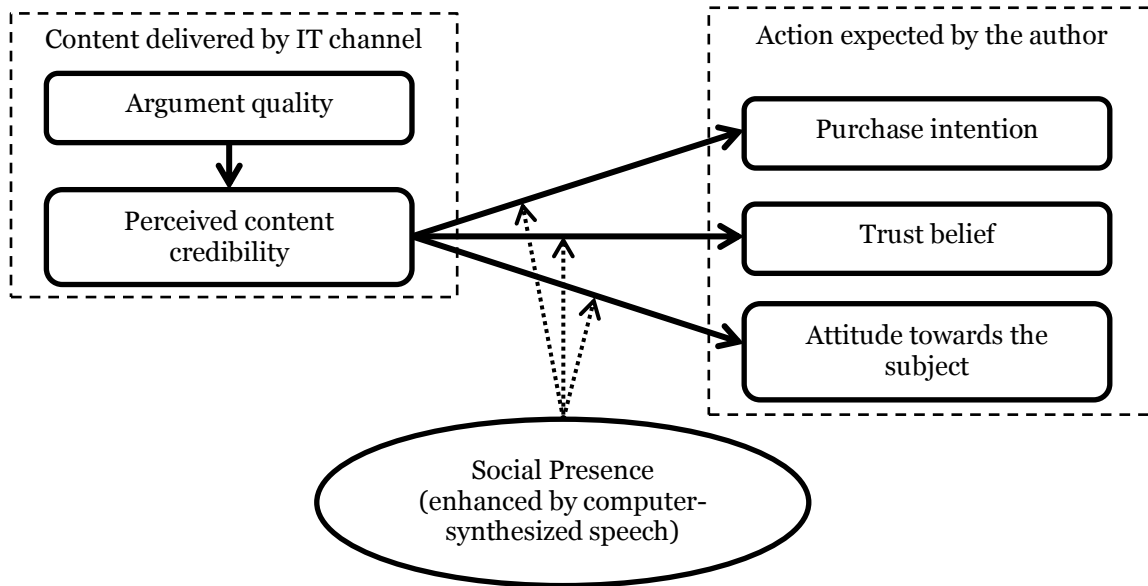
Empirical work finds that when people know that the message is delivered by a computer, there is no difference in perception of speech – the effect of human voice on persuasiveness is no greater than the effect of synthesized voice (Stern et al. 2006) and in some cases of problem solving situations, computer speech is actually more persuasive (Burgoon et al. 2000). Some researchers claimed also that using text synthesis to deliver a message instead of text may increase the credibility of the message source (Burgoon et al. 2000; Nass and Lee 2001), but empirical testing of this hypothesis awaits. The online reviews literature shows that source credibility has a strong effect on review's credibility, defined as the recipient's belief in message reliability and it contributes to the review's persuasiveness (Cheung et al. 2012; Smith et al. 2005).

The limited work on the role of computer-synthesized speech in computer-mediated, peer-to-peer commercial communication (Qiu and Benbasat 2005; Wang et al. 2007), confirms that even if a computer generates the voice message, speech is more enjoyable than a text message. Receivers of such a message focus more on the content of the message than they do when they experience text and speech or text only treatments (Qiu and Benbasat 2005). Furthermore, presentation of information via computer-synthesized speech leads to better long-term recognition of the information (Gathercole and Conway 1988). These results are important in the context of persuasion because increased attention helps individuals to focus on argument quality, which has stronger and longer effect on recipients (Petty and Cacioppo 1986). Moreover, because the receiver perceives speech as coming from someone more akin to them, rather than a machine, persuasiveness may increase further (Burger et al. 2004; Nass and Lee 2001).

## **Hypotheses and Research Model**

Building on the above literature and the definition of online content persuasiveness, we propose the following hypotheses and research model. We focus on the argument quality and perceived credibility as the elements of the online reviews content. To measure the persuasiveness, we recognize the most common motivations for writing online reviews (Yoo and Gretzel 2008) and existing research about online reviews persuasion (Schlosser et al. 2006; Sparks et al. 2013), and we use purchase intention, trust belief and attitude towards the subject as proxies of the expected action.

- H1. Argument quality of the structured online review, including claim, backing, and data increases perceived review credibility.
- H2. Higher perceived review credibility increases purchase intention.
- H3. Higher perceived review credibility increases trust belief.
- H4. Higher perceived review credibility increases attitude towards the subject.
- H5. Computer-synthesized speech as a message delivery method strengthens the effect of the review credibility on purchase intention.
- H6. Computer-synthesized speech as a message delivery method strengthens the effect of the review credibility on trust belief.
- H7. Computer-synthesized speech as a message delivery method strengthens the effect of the review credibility on attitude towards the subject.



**Figure 1. Research model**

## Method

We use a set of lab and field experiments to test the hypotheses. Both studies will be conducted in English and Italian for replication across different languages and to ensure that the effect of technology persuasiveness is language independent. Because the study is based on languages and speech, participants recruited for all studies will be native speakers to avoid confounds related to message understanding.

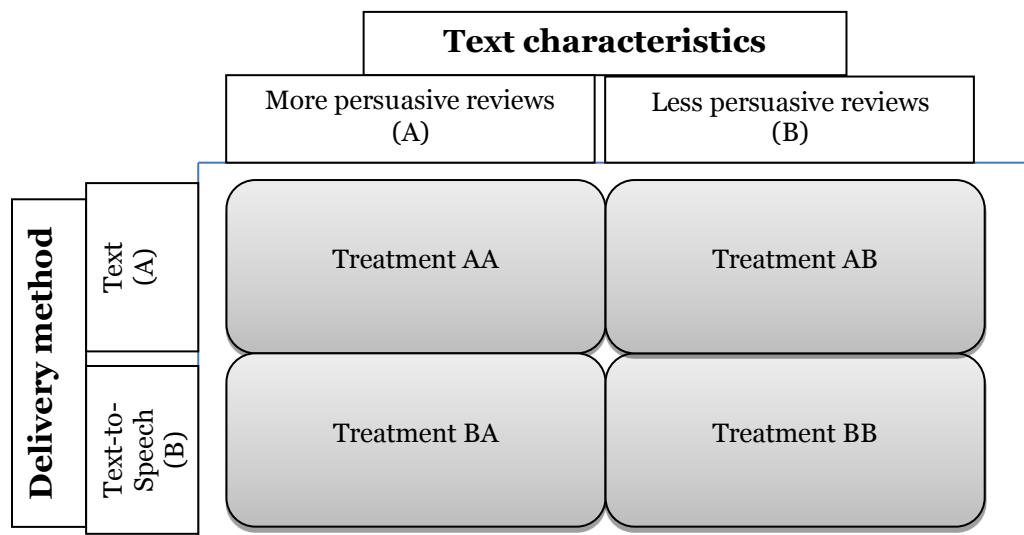
To increase experimental control we design artificial online reviews using a sample of real reviews. We use R's 'tm' package (Feinerer et al. 2008) to extract two sets of frequent words from 4047 English and 861 Italian reviews of a large European hotel. All the reviews used in the study were posted on a popular online review sites during the April 2011 to April 2014 timeframe. The artificial reviews will be standardized in terms of correctness and style. To provide the speech synthesis of online reviews to the participants, for both the English and Italian language sub-groups. We will use a male voice of high quality text-to-speech (TTS) software.

### Study 1

We will use a 2x2 factorial design (see Figure 2) to test our hypotheses. The participants will be randomly assigned to four groups. In a lab, each of them will see exactly the same fictional hotel recommendation website with peer reviews delivered by traditional text or by TTS software. We design the first factor as delivery method – some participants will read text reviews and some will listen to the reviews delivered by TTS. The second factor – text characteristics – also has two levels: 'standard' reviews and 'persuasive' reviews. We will create both sets using similar terms and the same subject matter. The difference between the treatments will pertain to only the composition of the reviews. We will construct persuasive reviews following Toulmin's model and other practices based on it.

Participants will use headphones to provide clear sound of computer-synthesized speech and to separate them from surrounding noise and distractions. To ensure that all groups are in similar conditions, participants will have to click on a title of a review to display text or reproduce sound. After having time to read or listen to the reviews each individual will complete a survey about perceived credibility, perceived usefulness and perceived persuasiveness of the reviews. We will collect also information about experience with online reviews and readiness to use speech-based online reviews in the future. This study will help to examine how the structure of a review changes the perception of peer-generated message credibility and

persuasiveness, as well as it will show if there is any difference in perception of speech-delivered online reviews and text reviews.



**Figure 2. Experimental design**

### **Study 2 – Field Experiment**

In the second study, we will examine the impact of a message delivery method of the reviews on the recipients' behavior, by evaluating actual decisions by experiment's participants. It will be a field experiment, using a factorial design, similar to the first study. Participants recruited through Amazon Mechanical Turk will be redirected to a fictional website with artificially created hotel reviews. Both factors in the design of the field study are analog to Study 1. Participants will choose a preferred hotel on the basis of text or speech reviews. To increase the realism of experimental conditions, participants will have a chance to win a short stay in the chosen hotel (all hotels will be in the same city to avoid travel considerations skewing results). We will collect participants' final choice; all their activities and time they need to complete the task. Comparison of all groups will give results showing if and how strong is the persuasive effect of computer-synthesized speech, and what is the reaction of participants to each condition in realistic environment. Additionally, each of four groups will see reviews of a control hotel and a treatment hotel, which will help to measure the difference between persuasiveness of the argument and persuasiveness of the technology.

### **Conclusion**

This paper presents a research-in-progress study about persuasiveness of computer-synthesized speech in the e-commerce context. Building on the results of planned experiments, we want to examine how the argument quality of the message affects the persuasiveness of online reviews and how computer-synthesized speech affects the behavior of online reviews readers. We expect that the results of the research will help both theorists and practitioners. It will improve understanding of individuals' reaction to different delivery technology of online reviews and commercial messages. It will be also useful for managers, as it will evaluate a new way to present text reviews in online recommendation systems. Companies which base their business model on online reviews (e.g. recommendation systems, booking engines, online shops) will receive insight into a new instrument to improve their services and to make them more appealing to consumers. They will be able to adjust themselves to match more with people's expectations.

## REFERENCES

- Bajarin, B. 2011. "In the Future, We Will All Talk to Computers," *Time* (available at <http://techland.time.com/2011/10/17/in-the-future-we-will-all-talk-to-computers/>).
- Benbasat, I. 2010. "HCI Research: Future Challenges and Directions," *AIS Transactions on Human-Computer Interaction* (2:2), pp. 16–21.
- Bhattacharjee, A., and Sanford, C. 2006. "Influence Processes for Information Technology Acceptance: An Elaboration Likelihood Model," *MIS Quarterly* (30:4), pp. 805–825.
- Burger, J. M., Messian, N., Patel, S., Prado, A. del, and Anderson, C. 2004. "What a Coincidence! The Effects of Incidental Similarity on Compliance," *Personality and Social Psychology Bulletin* (30:1), pp. 35–43.
- Burgoon, J. K., Bonito, J. A., Bengtsson, B., Cederberg, C., Lundeberg, M., and Allspach, L. 2000. "Interactivity in human–computer interaction: a study of credibility, understanding, and influence," *Computers in Human Behavior* (16:6), pp. 553–574.
- Cheung, C. M.-Y., Sia, C.-L., and Kuan, K. K. Y. 2012. "Is This Review Believable? A Study of Factors Affecting the Credibility of Online Consumer Reviews from an ELM Perspective," *Journal of Association of Information Systems* (13:8), pp. 618–635.
- Cialdini, R. B., and Trost, M. R. 1998. "Social influence: Social norms, conformity and compliance.," (available at <http://psycnet.apa.org/psycinfo/1998-07091-021>).
- Duan, W., Gu, B., and Whinston, A. B. 2008. "Do online reviews matter?—An empirical investigation of panel data," *Decision support systems* (45:4), pp. 1007–1016.
- Feinerer, I., Hornik, K., and Meyer, D. 2008. "Text Mining Infrastructure in R," *Journal of Statistical Software* (25:5).
- Fogg, B. J., and Tseng, H. 1999. "The Elements of Computer Credibility," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '99*, New York, NY, USA: ACM, pp. 80–87.
- Gathercole, S. E., and Conway, M. A. 1988. "Exploring long-term modality effects: Vocalization leads to best retention," *Memory & Cognition* (16:2), pp. 110–119.
- "Global Trust in Advertising and Brand Messages." 2013., New York: The Nielsen Company.
- Goldenberg, J., Libai, B., and Muller, E. 2001. "Talk of the network: A complex systems look at the underlying process of word-of-mouth," *Marketing letters* (12:3), pp. 211–223.
- Hong, S., and Park, H. S. 2012. "Computer-mediated persuasion in online reviews: Statistical versus narrative evidence," *Computers in Human Behavior* (28:3), pp. 906–919.
- Joo, Y. K., and Lee, J.-E. R. 2014. "Can 'The Voices in the Car' Persuade Drivers to Go Green?: Effects of Benefit Appeals from In-Vehicle Voice Agents and the Role of Drivers' Affective States on Eco-Driving," *Cyberpsychology, Behavior, and Social Networking* (17:4), pp. 255–261.
- Kim, D., and Benbasat, I. 2006. "The Effects of Trust-Assuring Arguments on Consumer Trust in Internet Stores: Application of Toulmin's Model of Argumentation," *Information Systems Research* (17:3), pp. 286–300.
- Kim, D., and Benbasat, I. 2009. "Trust-Assuring Arguments in B2C E-commerce: Impact of Content, Source, and Price on Trust," *Journal of Management Information Systems* (26:3), pp. 175–206.
- Lee, K. M., and Nass, C. 2003. "Designing Social Presence of Social Actors in Human Computer Interaction," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '03*, New York, NY, USA: ACM, pp. 289–296.
- Mattingly, I. G. 1970. "Speech synthesis for phonetic and phonological models," *Haskins Labs. Stat. Rep. Speech* (23), pp. 117–149.
- McGuire, K. 2013. "How Business Travelers Buy: Hotel Pricing in a Social World," *The Analytic Hospitality Executive*, October 24 (available at <http://blogs.sas.com/content/hospitality/2014/09/02/businesstravelerssocialworld/>; retrieved February 5, 2015).
- McGuire, K. 2014. "Pricing in a Social World: How consumers use ratings, reviews and price when choosing a hotel," *The Analytic Hospitality Executive*, September 2 (available at <http://blogs.sas.com/content/hospitality/2013/10/24/pricing-in-a-social-world2/>; retrieved February 5, 2015).
- Mudambi, S. M., and Schuff, D. 2010. "What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.Com," *MIS Q.* (34:1), pp. 185–200.
- Nass, C., and Gong, L. 2000. "Speech Interfaces from an Evolutionary Perspective," *Commun. ACM* (43:9), pp. 36–43.

- Nass, C., and Lee, K. M. 2001. "Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction," *Journal of Experimental Psychology: Applied* (7:3), pp. 171–181.
- Nass, C., Steuer, J., and Tauber, E. R. 1994. "Computers are social actors," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, pp. 72–78 (available at <http://dl.acm.org/citation.cfm?id=191703>).
- O’Keefe, D. J. 2004. "Trends and prospects in persuasion theory and research," in *Perspectives on persuasion, social influence, and compliance* J. S. Seiter and R. H. Gass (eds.), Boston: Pearson/Allyn and Bacon, pp. 31–43.
- Park, D.-H., Lee, J., and Han, I. 2007. "The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement," *International Journal of Electronic Commerce* (11:4), pp. 125–148.
- Pavlou, P. A., and Dimoka, A. 2006. "The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation," *Information Systems Research* (17:4), pp. 392–414.
- Petty, R. E., and Cacioppo, J. T. 1986. "The Elaboration Likelihood Model of Persuasion," in *Communication and Persuasion* Springer Series in Social Psychology, Springer New York, pp. 1–24 (available at [http://link.springer.com/chapter/10.1007/978-1-4612-4964-1\\_1](http://link.springer.com/chapter/10.1007/978-1-4612-4964-1_1)).
- Qiu, L., and Benbasat, I. 2005. "An Investigation into the Effects of Text-To-Speech Voice and 3D Avatars on the Perception of Presence and Flow of Live Help in Electronic Commerce," *ACM Trans. Comput.-Hum. Interact.* (12:4), pp. 329–355.
- Racherla, P., Mandviwalla, M., and Connolly, D. J. 2012. "Factors affecting consumers’ trust in online product reviews," *Journal of Consumer Behaviour* (11:2), pp. 94–104.
- Schlosser, A. E., White, T. B., and Lloyd, S. M. 2006. "Converting Web Site Visitors into Buyers: How Web Site Investment Increases Consumer Trusting Beliefs and Online Purchase Intentions," *Journal of Marketing* (70:2), pp. 133–148.
- Smith, D., Menon, S., and Sivakumar, K. 2005. "Online peer and editorial recommendations, trust, and choice in virtual markets," *Journal of Interactive Marketing* (19:3), pp. 15–37.
- Sparks, B. A., Perkins, H. E., and Buckley, R. 2013. "Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior," *Tourism Management* (39), pp. 1–9.
- Stern, S. E., Mullennix, J. W., and Yaroslavsky, I. 2006. "Persuasion and social perception of human vs. synthetic voice across person as source and computer as source conditions," *International Journal of Human-Computer Studies* (64:1), pp. 43–52.
- Torning, K., and Oinas-Kukkonen, H. 2009. "Persuasive System Design: State of the Art and Future Directions," in *Proceedings of the 4th International Conference on Persuasive Technology* Persuasive ’09, New York, NY, USA: ACM, pp. 30:1–30:8.
- Toulmin, S. E. 2003. *The uses of argument*, Cambridge University Press.
- Wang, L. C., Baker, J., Wagner, J. A., and Wakefield, K. 2007. "Can a retail web site be social?," *Journal of marketing* (71:3), pp. 143–157.
- Yin, D., Bond, S., and Zhang, H. 2014. "Anxious or Angry? Effects of Discrete Emotions on the Perceived Helpfulness of Online Reviews," *Management Information Systems Quarterly* (38:2), pp. 539–560.
- Yoo, K. H., and Gretzel, U. 2008. "What Motivates Consumers to Write Online Travel Reviews?," *Information Technology & Tourism* (10:4), pp. 283–295.

## APPENDIX 1

The table below contains two examples of the reviews prepared for the experiment. All reviews in the experiment are based on real reviews posted on a prominent online review site. They were created using the same subject matter and keeping similar length.

The review on the right follows the Toulmin's model and includes claim, data and backing. The review on the left is a control review and it has only claims. To ensure standardization of the topic and length of the text, there is more than only one claim in the control reviews.

To make it easier to identify each element of Toulmin's model in the table, claims are bold, data is underlined and backing is in italics.

Review based on Toulmin's model	Control review
<p><b>It is a good centrally located, and old fashioned hotel with very good service</b>  <u>- they try everything to please</u> <i>including complimentary non-alcoholic mini bar, free Wi Fi, tea and coffee and biscuits in the lobby, large buffet breakfast, comfortable beds and good pillow choices with clean rooms.</i></p>	<p><b>This is a good hotel located close to the city's Central Station. The room was clean, the service friendly and they offered free internet. The hotel had a small area in which you could get free service for coffee and tea with some biscuits. Breakfast itself was pretty big.</b></p>

**Table 1. Examples of experimental reviews**