Mining a corpus of online hotel reviews: a pilot study

Phoebe Lin and Amy Suen (The Hong Kong Polytechnic University, Hong Kong)

Electronic word-of-mouth (eWOM) is vital in today's competitive hospitality market. It is even more important to luxury hotels, which set hospitality standards worldwide. Online reviews on *TripAdvisor*, for example, are influential because they have a large readership. Studies have also shown their immediate impact on booking intentions (Vermeulen & Seegers, 2009; Ye et al., 2011).

This paper reports the findings of a corpus-driven study which examined 20,395 reviews about 12 Hong Kong 5-star hotels on *TripAdvisor*. The study aims to address two research questions about how the 5-star hotels reply to reviews while striving to promote their luxury brand image. These two research questions are: 1) how often do hotels reply to reviews? and 2) how do hotels reply to positive and negative reviews? Data analysis consists of two stages: quantitative analyses of reply counts and computer-assisted analyses of the language of the replies.

The results indicate that hotels reply more often to negative than positive reviews (reply rates of 65% vs 45%). There has also been a radical change in 'reply culture' over the past 12 years (2004-2016). In 2008, no hotel replied to reviews. However, in 2016, hotels replied to almost every review posted on TripAdvisor.

These replies, totalling 56,000 words, were then compiled into two corpora (replies to positive vs negative reviews). Keyness analyses were then applied (using Wmatrix) to reveal lexical, grammatical and semantic features that distinguish replies to negative reviews from replies to positive reviews. Amongst other findings, we noted that replies to low rating reviews are characterised by a tone of regret, ownership and reassurance. However, replies to high rating reviews are characterised by an affective and informal tone, as well as reiterations of the hotels' competitive advantages.

This study provides a corpus-driven description of the ways in which 5-star hotels respond to online reviews. It highlights the various linguistic resources that are often exploited in the online review environment. These findings may shed light on future research on communication strategies in tourism and hospitality management, and inform the development of relevant ESP teaching materials.

References

Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management, 30*, 123-127.

Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveller behaviour: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behaviour, 27*(2), 634-639.