

# Chapter 81

## Words that Fascinate the Listener: Predicting Affective Ratings of On-Line Lectures

**Felix Weninger**

*Technische Universität München, Germany*

**Pascal Staudt**

*Technische Universität München, Germany*

**Björn Schuller**

*Technische Universität München, Germany*

### ABSTRACT

*In a large scale study on 843 transcripts of Technology, Entertainment and Design (TED) talks, the authors address the relation between word usage and categorical affective ratings of lectures by a large group of internet users. Users rated the lectures by assigning one or more predefined tags which relate to the affective state evoked in the audience (e. g., ‘fascinating’, ‘funny’, ‘courageous’, ‘unconvincing’ or ‘long-winded’). By automatic classification experiments, they demonstrate the usefulness of linguistic features for predicting these subjective ratings. Extensive test runs are conducted to assess the influence of the classifier and feature selection, and individual linguistic features are evaluated with respect to their discriminative power. In the result, classification whether the frequency of a given tag is higher than on average can be performed most robustly for tags associated with positive valence, reaching up to 80.7% accuracy on unseen test data.*

### INTRODUCTION

Sensing affect related states, including interest, confusion, or frustration, and adapting behavior accordingly, is one of the key capabilities of hu-

mans; consequently, simulating such abilities in technical systems through signal processing and machine learning techniques is believed to improve human-computer interaction in general (Schuller & Weninger, 2012) and computer based learning

DOI: 10.4018/978-1-4666-6042-7.ch081

in particular (Aist, Kort, Reilly, Mostow, & Picard, 2002; Forbes-Riley & Litman, 2010). Important abilities of affective tutors or lecturers, besides emotional expressivity (Huang, Kuo, Chang, & Heh, 2004), include the choice of appropriate wording, which has been found to be highly important in computer based tutoring to support the learning outcome (Narciss & Huth, 2004). Furthermore, there is increased evidence for the influence of affect related states on the learning process (Craig, Graesser, Sullins, & Gholson, 2004; Bhatt, Evens, & Argamon, 2004; Forbes-Riley & Litman, 2007). In particular, previous studies highlighted the relation between system responses in a tutoring dialogue and student affect (Pour, Hussein, Al Zoubi, D'Mello, & Calvo, 2011); it turned out, for example, that dialogue acts of an automated tutor influence student uncertainty (Forbes-Riley & Litman, 2011). However, these studies do not take into account the linguistic content of lectures as a whole; hence, we aim to bridge this gap by addressing the automatic assignment of categorical affective ratings by a large audience to on-line lectures from the TED talks website ([www.ted.com/talks](http://www.ted.com/talks)). This prediction is based on learning the relation between linguistic features of the speech transcripts and the ratings given by the audience, which comprises many thousands of internet users in our case. Such automatic predictions can be immediately useful to evaluate the quality of lectures given by a distant education system, and to gain insight into which lecture topics or lecturing strategies are related to certain affective states. The aspect of predicting the induced affect from the lecturers' speech has—to our knowledge—not been addressed in a systematic fashion so far: Rather, in (Forbes-Riley & Litman, 2011), features from student responses to the system and abstract goals of the dialogue manager are used to analyze student affect. In this respect, that study is somewhat related to sentiment analysis (Schuller & Knaup, 2010) or opinion mining (Turney, 2002), where the goal is to deduce the affect of the users from

written reviews. However, in our study we aim at predicting the users' affective ratings based on the lectures themselves. This also distinguishes our contribution from the large body of literature on prediction of (ordinal-scale) movie ratings—for a recent study on the public Internet Movie Database (IMDB), we refer to (Marovic, Mihokovic, Miksa, Pribil, & Tus, 2011). In that field, in contrast to our study, the vast majority of approaches seem to be exploiting similarities in user profiles rather than features of the rated objects (*instances* in terms of machine learning), such as in (Marlin, 2003).

Finally, in contrast to many previous studies focusing on singular affects or ratings on a single 'good or bad' scale, we investigate the multi-label categorical ratings from the TED talks website which are given by internet users through assignment of tags to each talk. The tag set is determined by the creators of the TED talks website. These tags are on the one hand directly associated with the emotion evoked in the audience (e.g., 'obnoxious', 'funny'); on the other hand, they can refer to perceived attributes of the speaker resulting in a certain affect of the audience (e.g., 'courageous', which may result in 'feeling moved'). Third, the tags may describe the argumentative structure of the talk (e.g., 'long-winded', 'unconvincing'), which is arguably reflected in affective states such as 'boredom' or 'confusion' which are often investigated in the context of tutoring (Pour et al., 2011; Forbes-Riley & Litman, 2011). For the most part, these tags refer to emotion-related states (e.g., confusion, feeling inspired) rather than full blown 'Big 6' emotions, while 'obnoxious' (arguably, the 'strongest' of the 14 tags) could roughly correspond to disgust. Furthermore, in most cases, tags can be classified into those inducing positive or negative valence (e.g., 'confusing' and 'long-winded' for negative valence, and 'inspiring' or 'fascinating' for positive valence).

On lectures from the TED talks website, a slightly tongue-in-cheek analysis has been performed as to the relation of linguistic content and the user ratings ([www.get-tedpad.com](http://www.get-tedpad.com)). This

11 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:  
[www.igi-global.com/chapter/words-that-fascinate-the-listener/108797](http://www.igi-global.com/chapter/words-that-fascinate-the-listener/108797)

---

## Related Content

### Hidden Markov Model Based Visemes Recognition, Part I: AdaBoost Approach

Say Wei Foo and Liang Donga (2009). *Visual Speech Recognition: Lip Segmentation and Mapping* (pp. 326-355).

[www.irma-international.org/chapter/hidden-markov-model-based-visemes/31073](http://www.irma-international.org/chapter/hidden-markov-model-based-visemes/31073)

### Robustness Analysis of Patchwork Watermarking Schemes

Hyoung-Joong Kim, Shijun Xiang and In-Kwon Yeo (2008). *Digital Audio Watermarking Techniques and Technologies: Applications and Benchmarks* (pp. 172-186).

[www.irma-international.org/chapter/robustness-analysis-patchwork-watermarking-schemes/8331](http://www.irma-international.org/chapter/robustness-analysis-patchwork-watermarking-schemes/8331)

### Itakura-Saito Nonnegative Factorizations of the Power Spectrogram for Music Signal Decomposition

Cédric Févotte (2011). *Machine Audition: Principles, Algorithms and Systems* (pp. 266-296).

[www.irma-international.org/chapter/itakura-saito-nonnegative-factorizations-power/45489](http://www.irma-international.org/chapter/itakura-saito-nonnegative-factorizations-power/45489)

## Conclusion

(2020). *Grammatical and Syntactical Approaches in Architecture: Emerging Research and Opportunities* (pp. 324-334).

[www.irma-international.org/chapter/conclusion/245867](http://www.irma-international.org/chapter/conclusion/245867)

### Advances in Computer Speech Synthesis and Implications for Assistive Technology

H. Timothy Bunnell and Christopher A. Pennington (2010). *Computer Synthesized Speech Technologies: Tools for Aiding Impairment* (pp. 71-91).

[www.irma-international.org/chapter/advances-computer-speech-synthesis-implications/40859](http://www.irma-international.org/chapter/advances-computer-speech-synthesis-implications/40859)