

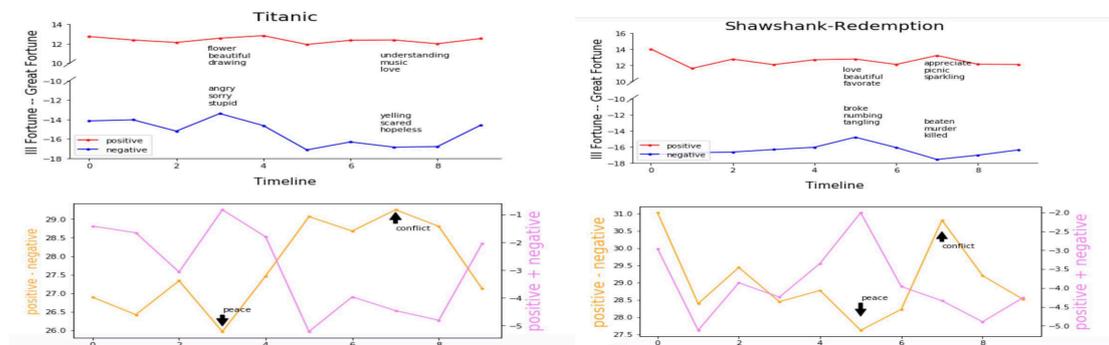
# The Shape of Stories

Huimin Xu<sup>1</sup>, Cheng-Jun Wang<sup>1\*</sup>

(1) School of Journalism and Communication, Nanjing University, Nanjing 210093, China, 18362929216@163.com

Humans are animals of storytelling, which is always a collective process. Movies, books, TV shows, across all kinds of culture products socially created and publicized, only those narratives who successfully structured the emotions of readers will be remembered (Vonnegut, 1981).

To understand what kind of stories are more likely to get attention and become part of collective memory, we analyzed the scripts and reviews of 16,364 movies collected from imdb.com. We first construct a vector representing “fortunate” by retrieving the pre-trained embedding vectors (Pennington et al., 2014) of two groups of anchoring words, including “failure”, “sad”, “unlucky” et al. for negative emotion and “success”, “happy”, and “lucky” for positive emotion. We then subtract the averages vector of the first group of words from the average of the second group to create the “fortunate” vector. Then we sliced the movie scripts into ten chapters. For each chapter, we calculated the averaged cosine distance between the ‘fortunate’ vector and every word in the chapter. We find this analysis successfully captures the dramatic shifts of narrative, as shown in Figure 1.



**Figure 1.** The shape of *Titanic* (left panel) and *Shawshank's Redemption* (right panel).

By separating positive and negative words and measuring the cosine distance of these two groups of word to the anchoring “fortunate” vector separately, we obtained two lines capturing the positive (red) and negative (blue) emotion of stories. Summing up the positive and negative values gives the pink line in the lower panel, which represents the emotional bias. Subtracting negative from positive values gives the orange line in the lower panel, which represents the emotional intensity. Dramatic shift typically happens when the narrative is intensive and strongly biased. Examples include the seventh part of *Titanic*, when the ship tragically sinking yet Jack and Rose were saying sweet words to each other. Also, in *Shawshank's Redemption*, when Jack successfully escaped at a rainy night, the plot comes to the most intensive and positively biased part.

To associate stories shapes with their success or failure, we used an artificial neural network to predict rates from the vectors featuring the shapes. We found that successful movies typically involve more dramatic shifts and are more biased towards positive emotion in general. More analysis will be done given an opportunity to present.

### **References**

Pennington, J., Socher, R., and Manning, C. D. Glove: Global vectors for word representation. In Empirical Methods in Natural Language Processing (EMNLP) , pp. 1532–1543, 2014.

Vonnegut K. (1981). Palm Sunday. RosettaBooks LLC, New York.