Automatic Detection of Memorable Spoken Quotes *

Fajri Koto† (NAIST/UI), ◎Sakriani Sakti, Graham Neubig, Tomoki Toda‡ (NAIST),
Mirna Adriani (UI), Satoshi Nakamura (NAIST)

1 Introduction

Social robots are gaining increasing interest, and therefore understanding the ways humans express themselves and engage their listeners is becoming more important. Here, we focus on studying natural expressiveness and its effects during public speeches. Through history, the best speeches of all time normally feature memorable quotes that genuinely inspire the audience. For instance, John F. Kennedy's famous quote: '*Ask not what your country can do, ask what you can do for your country*', has inspired many generations since he gave this speech in 1961[1]. While some public speeches may have inspired many individuals, they raise deeper questions. Why can some public speeches be memorable and inspirational for audience, while some others can not? In this study, we attempt to answer these questions by developing a method for automatic detection of memorable spoken quotes of real public speeches.

2 Memorable Spoken Quotes Detection

Most techniques developed so far for memorable quote detection have focused primarily on the processing of text [1, 2, 3], we are interested in discovering memorable spoken quotes of real public speeches.

2.1 Corpus Construction

Many famous people have given speeches on TED[2] and inspired people by their memorable words. TED also provided 'TED Quotes,' which collects memorable quotes from TED talks, and allow people to share their favorite quotes. The most popular quotes can have more than a thousand shares. The collected memorable quotes resulted in a total of 2118 speech transcription segment files. Then we randomly selected 2118 transcriptions from other TED segment files and labeled them as non-memorable quotes.

2.2 Features of Spoken Quotes

Bandersky et al. defined three kinds of linguistic features: lexical, punctuation, and part-of-speech (POS) features which have been useful for memorable quote detection [1]. As we focus on spoken utterance of memorable quotes, we utilize lexical and part-of-speech features, but exclude punctuation features. In addition, we included hasSynonym and hasAntonym features in our experiment. A detailed descriptions are shown in Table 1.

2.3 Classifier

We investigated three classifiers: Neural Networks (NN), Naive Bayes (NB), and Support Vector Machines (SVM). We also performed feature selection, greedily selecting and adding the most effective feature one at a time.

3 Experimental Evaluation

3.1 Memorable Quote Detection

Table 2 shows the performance of all classifiers (with 5-fold cross validation) after feature selection. As comparison, we also include the performance of the classifier using the features proposed by Bandersky (denoted as "Baseline"). The results reveal that our proposed features give better accuracy than the baseline, and the best results were achieved by the use of acoustic features.

Next, we combine all selected features from all classifier into one union set of selected features. As there is some overlap of features, we finally have 12 linguistic features and 9 acoustic features in total. The result shows the accuracy of memorable quote detection based on SVM classifier, using: (1) 12 selected linguistic features only with 66.45% accuracy,

*記憶に残る通用引用の自動検出, Koto Fajri(NAIST/UI), Sakti Sakriani, Neubig Graham, 戸田智基 (NAIST), Adriani Mirna (UI), 中村哲 (NAIST)

†Now at Samsung Research, Indonesia

‡Now at Nagoya University

[1]http://www.ushistory.org/

[2]http://www.ted.com/

Table 1: Linguistic feature sets for a particular quote s.

Feature	Description
	Lexical
#capital	Number of capitalized words in s
#quantifier	Number of universal quantifiers in s
#stops	Number of common stopwords in s
beginStop	True if s begins with a stopword, False otherwise
hasDialog	True if s contains at least one of say, says, said
#abstract	Number of abstract concepts (e.g., adventure, charity, stupidity) in s
	Part of Speech
#POS	POS = noun, verb, adjective, adverb, pronoun
hasComp	True if s contains a comparative adjective or adverb, False otherwise
hasSuper	True if s contains a superlative adjective or adverb, False otherwise
hasPP	True if s contains a verb in past participle, False otherwise
hasSynonym	True if s contains two words that are synonymous, False otherwise
hasAntonym	True if s contains two words are antonyms of each other, False otherwise
#IGSeq[i]	Count of the POS sequence with i-th highest information gain in s

(2) 9 selected acoustic features only with 68.06 and (3) combination of the selected linguistic and acoustic features with the highest, 70.4% accuracy. The results reveal that the classifier with combination features perform better than the classifier with linguistic or acoustic features only.

Table 2: Accuracy of memorable quote detection (the chance rate is 50.0%).

	Baseline	Proposed	
Classifier	Linguistic	Linguistic	Acoustic
NN	63.98%	64.87%	**67.71**%
NB	62.91%	65.04%	**68.18**%
SVM	64.80%	66.71%	**68.08**%

Table 3: POS-tag sequences selected for memorableness analysis (MQ = Memorable Quotes and NM = Non-Memorable Quotes)

Sequence	Example	#MQ	#NM
CC-PRP-VBD	'but i thought', 'and i introduced'	43	124
NN-VBZ-DT	'belief is the', 'education is a'	155	45
JJ-NN-NN	'national automobile slum', 'quiet screaming desperation'	236	183
PRP-VBZ-IN	'it is as', 'it is like'	95	39
NN-VBZ-RB	'innovation is not', 'privacy is not'	165	50

3.2 Memorableness Analysis

We further analyze features selected by the feature selection procedure. For acoustic features, we found out that F0 had the highest weight. It indicates that the prosody of the utterance is a significant feature that distinguish between memorable quotes and non-memorable quotes.

For linguistic features, the selected features include beginStop, #noun, #adjective and some POS-tag sequences described in Table 3. The sentences containing CC-PRP-VBD sequence tend to be non-memorable quotes, which indicates that memorable quotes seldom use conjunctions following by personal pronouns that represent specific things or particular individuals. On the other hand, sentences with POS sequences of NN-VBZ-DT, JJ-NN-NN, PRP-VBZ-IN and NN-VBZ-RB tend to be memorable quotes. These POS sequences are mainly used for definition, elaboration and explanation types of sentences.

4 Conclusion

In this study, we discuss the possibilities to automatically detect the memorable spoken quotes of real public speeches based on linguistic and acoustic features. The results reveal that we can distinguish between memorable quotes and non-memorable quotes with up to 70.4% accuracy. Based on the analysis of the selected features, the results reveal that most memorable quotes have definition, elaboration and explanation type of sentences, and the prosody of the utterance is a significant acoustic feature that distinguishes between memorable quotes and non-memorable quotes. Further details of our study can be found in [4].

Acknowledgements Part of this work was supported by JSPS KAKENHI Grant Number 26870371.

References

[1] M. Bandersky et al., "A dictionary of wisdom and wit: Learning to extract quotable phrase," in *Proc. NAACHL-HLT*, 2012.

[2] O. Kolak et al., "Generating links by mining quotations," in *Proc. ACM conference on Hypertext and hypermedia*, 2008.

[3] C. Danescu-Niculescu-Mizil et al., "You had me at hello: How phrasing affects memorability," in *Proc. ACL*, 2012.

[4] F. Koto et al., "A study on natural expressive speech: Automatic memorable spoken quote detection," in *Proc. IWSDS*, 2015.