Automated Topical Component Extraction Using Neural Network Attention Scores from Source-based Essay Scoring

Haoran Zhang
Department of Computer Science
University of Pittsburgh
Pittsburgh, PA 15260
colinzhang@cs.pitt.edu

Diane Litman
Department of Computer Science & LRDC
University of Pittsburgh
Pittsburgh, PA 15260
litman@cs.pitt.edu

Abstract

While automated essay scoring (AES) can reliably grade essays at scale, automated writing evaluation (AWE) additionally provides formative feedback to guide essay revision. However, a neural AES typically does not provide useful feature representations for supporting AWE. This paper presents a method for linking AWE and neural AES, by extracting Topical Components (TCs) representing evidence from a source text using the intermediate output of attention layers. We evaluate performance using a feature-based AES requiring TCs. Results show that performance is comparable whether using automatically or manually constructed TCs for 1) representing essays as rubric-based features, 2) grading essays.

1 Introduction

Automated essay scoring (AES) systems reliably grade essays at scale, while automated writing evaluation (AWE) systems additionally provide formative feedback to guide revision. Although neural networks currently generate state-of-the-art AES results (Alikaniotis et al., 2016; Taghipour and Ng, 2016; Dong et al., 2017; Farag et al., 2018; Jin et al., 2018; Li et al., 2018; Tay et al., 2018; Zhang and Litman, 2018), non-neural AES create feature representations more easily usable by AWE (Roscoe et al., 2014; Foltz and Rosenstein, 2015; Crossley and McNamara, 2016; Woods et al., 2017; Madnani et al., 2018; Zhang et al., 2019). We believe that neural AES can also provide useful information for creating feature representations, e.g., by exploiting information in the intermediate layers.

Our work focuses on a particular source-based essay writing task called the response-to-text assessment (RTA) (Correnti et al., 2013). Recently, an RTA AWE system (Zhang et al., 2019) was built by extracting rubric-based features related to the use of Topical Components (TCs) in an essay. However, manual expert effort was first required to create the TCs. For each source, the TCs consist of a comprehensive list of topics related to evidence which include: 1) important words indicating the set of evidence topics in the source, and 2) phrases representing specific examples for each topic that students need to find and use in their essays.

To eliminate this expert effort, we propose a method for using the interpretable output of the attention layers of a neural AES for source-based essay writing, with the goal of extracting TCs. We evaluate this method by using the extracted TCs to support feature-based AES for two RTA source texts. Our results show that 1) the feature-based AES with TCs manually created by humans is matched by our neural method for generating TCs, and 2) the values of the rubric-based essay features based on automatic TCs are highly correlated with human Evidence scores.

2 Related Work

Three recent AWE systems have used non-neural AES to provide rubric-specific feedback. Woods et al. (2017) developed an influence estimation process that used a logistic regression AES to identify sentences needing feedback. Shibani et al. (2019) presented a web-based tool that provides formative feedback on rhetorical moves in writing. Zhang et al. (2019) used features created for a random forest AES to select feedback messages, although human effort was first needed to create TCs from a source text. We automatically extract TCs using neural AES, thereby eliminating this expert effort.

Others have also proposed methods for pre-processing source information external to an essay. Content importance models for AES predict the parts of a source text that students should include when writing a summary (Klebanov et al., 2018).
Methods for extracting important keywords or keyphrases also exist, both supervised (unlike our approach) (Meng et al., 2017; Mahata et al., 2018; Florescu and Jin, 2018) and unsupervised (Florescu and Caragea, 2017). Rahimi and Litman (2016) developed a TC extraction LDA model (Blei et al., 2003). While the LDA model considers all words equally, our model takes essay scores into account by using attention to represent word importance. Both the unsupervised keyword and LDA models will serve as baselines in our experiments.

In the computer vision area, attention cropped images have been used for further image classification or object detection (Cao et al., 2015; Yuxin et al., 2018; Ebrahimpour et al., 2019). In the NLP area, Lei et al. (2016) proposed to use a generator to find candidate rationale and these are passed through the encoder for prediction. Our work is similar in spirit to this type of work.

### 3 RTA Corpus and Prior AES Systems

The essays in our corpus were written by students in grades 4 to 8 in response to two RTA source texts (Correnti et al., 2013): \( RTA_{MV} \) (2970 essays) and \( RTA_{SP} \) (2076 essays). Table 1 shows an excerpt from \( RTA_{MV} \), the associated essay writing prompt, and a student essay. The bolding in the source indicates evidence examples that experts manually labeled as important for students to discuss (i.e., TC phrases). Evidence usage in each essay was manually scored on a scale of 1 to 4 (low to high). The distribution of Evidence scores is shown in Table 2. The essay in Table 1 received a score of 3, with the bolding indicating phrases semantically related to the TCs from the source text.

To date, two approaches to AES have been proposed for the RTA: \( AES_{rubric} \) and \( AES_{neural} \). To support the needs of AWE, \( AES_{rubric} \) (Zhang and Litman, 2017) used a traditional supervised learning framework where rubric-motivated features were extracted from every essay before model training - Number of Pieces of Evidence (NPE) \(^1\), Concentration (CON), Specificity (SPC) \(^2\), Word Count (WOC). The two aspects of TCs introduced in Section 1 (topic words, specific example phrases) were used during feature extraction.

Motivated by improving stand-alone AES performance (i.e., when an interpretable model was not needed for subsequent AWE), Zhang and Litman (2018) developed \( AES_{neural} \), a hierarchical neural model with the co-attention mechanism in the sentence level to capture the relationship between the essay and the source. Neither feature engineering nor TC creation were needed before training.

### 4 Attention-Based TC Extraction: \( TC_{attn} \)

In this section we propose a method for extracting TCs based on the \( AES_{neural} \) attention level outputs. Since the self-attention and co-attention mechanisms were designed to capture sentence and phrase importance, we hypothesize that the attention scores can help determine if a sentence or phrase is important. We consider a sentence important if the \( TC_{attn} \) score associated with it is greater than a threshold.

---

**Table 1:** A source excerpt for the \( RTA_{MV} \) prompt and an essay with score of 3.

| Prompt | \( RTA_{MV} \) | \( RTA_{SP} \) |
|--------|----------------|----------------|
| Score 1| 852            | 538            |
| Score 2| 1197           | 789            |
| Score 3| 616            | 512            |
| Score 4| 305            | 237            |
| Total  | 2970           | 2076           |

**Table 2:** The Evidence score distribution of RTA.

---

\(^1\)An integer feature based on the list of topic words for each topic.

\(^2\)A vector of integer values indicating the number of specific example phrases (semantically) mentioned in the essay per topic.
phrase has important source-related information.

To provide intuition, Table 3 shows examples sentences from the student essay in Table 1. Bolded are phrases with the highest self-attention score within the sentence. Italicics are specific example phrases that refer to the manually constructed TCs for the source. $Attn_{sent}$ is the text to essay attention score that measures which essay sentences have the closest meaning to a source sentence. $Attn_{phrase}$ is the self-attention score of the bolded phrase that measures phrase importance. A sentence with a high attention score tends to include at least one specific example phrase, and vice versa. The phrase with the highest attention score tends to include at least one specific example phrase if the sentence has a high attention score.

Based on these observations, we first extract the output of two layers from the neural network: 1) the $attn_{sent}$ of each sentence, and 2) the output of the convolutional layer as the representation of the phrase with the highest $attn_{phrase}$ in each sentence (denoted by $cmn_{phrase}$). We also extract the plain text of the phrase with the highest $attn_{phrase}$ in each sentence (denoted by $text_{phrase}$). Then, our $TC_{attn}$ method uses the extracted information in 3 main steps: 1) filtering out $text_{phrase}$ from sentences with low $attn_{sent}$, 2) clustering all remaining $text_{phrase}$ based on $cmn_{phrase}$, and 3) generating TCs from clusters.

The first filtering step keeps all $text_{phrase}$ where the original sentences have $attn_{sent}$ higher than a threshold. The intuition is that lower $attn_{sent}$ indicates less source-related information.

The second step clusters these $text_{phrase}$ based on their corresponding representations $cmn_{phrase}$. We use k-medoids to cluster $text_{phrase}$ into $M$ clusters, where $M$ is the number of topics in the source text. Then, for $text_{phrase}$ in each topic cluster, we use k-medoids to cluster them into $N$ clusters, where $N$ is the number of the specific example phrases we want to extract from each topic. The outputs of this step are $M \times N$ clusters.

The third step uses the topic and example clustering to extract TCs. As noted earlier, TCs include two parts: topic words, and specific example phrases. Since our method is data-driven and students introduce their vocabulary into the corpus, essay text is noisy. To make the TC output cleaner, we filter out words that are not in the source text. To obtain topic words, we combine all $text_{phrase}$ from each topic cluster to calculate the word frequency per topic. To make topics unique, we assign each word to the topic cluster in which it has the highest normalized word frequency. We then include the top $K_{topic}$ words based on their frequency in each topic cluster. To obtain example phrases, we combine all $text_{phrase}$ from each example cluster to calculate the word frequency per example, then include the top $K_{example}$ words based on their frequency in each example cluster.

### 5 Experimental Setup and Results

Figure 1 shows an overview of four TC extraction systems. PositionRank is not designed for TC extraction methods to be evaluated. $TC_{manual}$ (upper bound) uses a human expert to extract TCs from a source text. $TC_{attn}$ is our proposed method and automatically extracts TCs using both a source text and student essays. $TC_{lda}$ (Rahimi and Litman, 2016) (baseline) builds on LDA to extract TCs from student essays only, while $TC_{pr}$ (baseline) builds on PositionRank (Florescu and Caragea, 2017) to instead extract TCs from only the source text.
Component

TC

we put them into only one topic and remove all word to a higher dimension with word embedding.

(2018) for neural training as shown in Table 4. Ta-

(NPE and sum of SPC features)

3 Zhang and Litman, 2017 will perform compara-

tion itself.

Downstream AES using our proposed TC extraction method on the

TC and will be stronger than when using

TC or

TC

will perform worse when using

TC or

TC

pr

the human Evidence score and the feature values

H1) the

AESrubric

model for scoring Evidence (Zhang and Litman, 2017) will perform comparably when extracting features using either

TC

attn

or

TC

manual

and will perform worse when using

TC

lda

or

TC

pr

H2) the correlation between the human Evidence score and the feature values (NPE and sum of SPC features) will be comparable when extracted using

TCattn

and

TCmanual

and will be stronger than when using

TCllda

and

TCpr

The experiment for H1 tests the impact of using our proposed TC extraction method on the downstream

AESrubric

task, while the H2 experiment examines the impact on the essay representation itself.

Following Zhang and Litman (2017), we stratify essay corpora: 40% for training word embeddings and extracting TCs, 20% for selecting the best embedding and parameters, and 40% for testing. We use the hyper-parameters from Zhang and Litman (2018) for neural training as shown in Table 4. Table 5 shows all other parameters selected using the development set.

Results for H1. H1 is supported by the results in Table 6, which compares the Quadratic Weighted Kappa (QWK) between human and

AESrubric Evidence scores (values 1-4) when

AESrubric

uses

TCmanual

versus each of the automatic methods.

TCattn

always yields better performance, and even significantly better than

TClmanual

Results for H2. The results in Table 7 support

H2. TCattn outperforms the two automated base-

lines, and for NPE even yields stronger correlations than the manual TC method.

Qualitative Analysis. The manually-created topic words for

RTAMVP

represent 4 topics, which are “hospital”, “malaria”, “farming” and “school”4. Although Table 5 shows that the automated list has more topics for topic words and might have broken one topic into separate topics, a good automated list should have more topics related to the 4 topics above. We manually assign a topic for each of the topic words from the different automated methods. TCllda has 4 related topics out of 9 (44.44%),

TCllda

has 6 related topics out of 19 (31.58%), and

TCllda

has 10 related topics out of 16 (62.50%). Obviously,

TCllda

preserves more related topics than our baselines.

Moving to the second aspect of TCs (specific example phrases), Table 8 shows the first 10 specific example phrases for a manually-created category that introduces the changes made by the MVP project5. This category is a mixture of different topics because it talks about the “hospital”, “malaria”, “school”, and “farming” at the same time.

TCattn

has overlap with

TCmanual

on different topics. However, TCllda mainly talks about “hospital”, because the nature of the LDA model doesn’t allow mixing specific example phrases about different topics in one category. Unfortunately, TCap

---

Table 5: Parameters for different models.

Prompt | Component | Parameter | TCllda | TCllda | TCllda
---|---|---|---|---|---|
RTAMVP | Topic Words | Number of Topics | 9 | 19 | 16
| Number of Words | 30 | 20 | 25
| Example Phrases | Number of Topics | 20 | 18
| Number of Phrases | 15 | 20 | 15
| RTA<sub>Space</sub> | Topic Words | Number of Topics | 15 | 20 | 10
| Number of Words | 10 | 10 | 20
| Example Phrases | Number of Topics | 10 | 4 | 9
| Number of Phrases | 20 | 50 | 20

Table 6: The performance (QWK) of

AESrubric

using different TC extraction methods for feature creation. The numbers in the parentheses show the model numbers over which the current model performs significantly better ($p < 0.05$). The best results between automated methods in each row are in bold.

Prompt | Feature | TClmanual (1) | TCllda (2) | TCllda (3) | TCllda (4)
---|---|---|---|---|---|
RTAMVP | NPE | 0.598 | 0.588 | 0.567 | 0.679
| SPC (sum) | 0.625 | 0.574 | 0.533 | 0.598
| RTA<sub>Space</sub> | NPE | 0.484 | 0.513 | 0.494 | 0.625
| SPC (sum) | 0.601 | 0.574 | 0.533 | 0.598

Table 7: Pearson’s r comparing feature values computed using each TC extraction method with human (gold-standard) Evidence essay scores. All correlation values are significant ($p \leq 0.05$). The best results between automated methods in each row are in bold.

---

4All Specific Example Phrases generated by different models can be found in the Appendix A.2.

5All Specific Example Phrases generated by different models can be found in the Appendix A.1.

---

These features are extracted based on TCs.
Table 8: Specific example phrases for the $RT^A_{MV P}$ progress topic.

| $T_{Comp}$                     | $T_{lda}$                             | $T_{pr}$                           | $T_{attn}$          |
|--------------------------------|---------------------------------------|-----------------------------------|---------------------|
| medicine most common diseases  | water connected hospital generator electricity | millennium villages project        | electricity running water irrigation set |
| running water electricity      | patients afford                        | unpaved dirt road                  | farmers could afford bed electricity hospital |
| hospital generator electricity | patients probably                     | bar sauri primary school           | better fertilizer medicine enough also |
| bed nets used every sleeping site | share beds                             | future hannah                      | rooms packed patients |
| hunger crisis addressed fertilizer seeds | receive treatment                    | sauri primary school               | food fertilizer crops get supply |
| tools needed maintain food supply | doctors clinical                      | villages project                   | five net costs 5     |
| no school fees                 | doctors clinical                       | millennium development goals        | nets net bed free    |
| school attendance rate way up  | water fertilizer knowledge             | village leaders                    | running water supplies schools almost |
| kids go school now             | receive treatment                      | dirt road                          |                     |
|                                | ...                                   | ...                                | ...                 |

does not include any overlapped specific phrase in the first 10 items; they all refer to some general example phrases from the beginning of the source article. Although there are some related specific example phrases in the full list, they are mainly about school. This is because the PositionRank algorithm tends to assign higher scores to words that appear early in the text.

6 Conclusion and Future Work

This paper proposes $TC_{attn}$, a method for using the attention scores in a neural AES model to automatically extract the Topical Components of a source text. Evaluations show the potential of $TC_{attn}$ for eliminating expert effort without degrading $AES_{rubric}$ performance or the feature representations themselves. $TC_{attn}$ outperforms baselines and generates comparable or even better results than a manual approach.

Although $TC_{attn}$ outperforms all baselines and requires no human effort on TC extraction, annotation of essay evidence scores is still needed. This leads to an interesting future investigation direction, which is training the $AES_{neural}$ using the gold standard that can be extracted automatically.

One of our next steps is to investigate the impact of TC extraction methods on a corresponding AWE system (Zhang et al., 2019), which uses the feature values produced by $AES_{rubric}$ to generate formative feedback to guide essay revision.

Currently, the $TC_{lda}$ are trained on student essays, while the $TC_{pr}$ only works on the source article. However, $TC_{attn}$ uses both student essays and the source article for TC generation. It might be hard to say that the superior performance of $TC_{attn}$ is due to the neural architecture and attention scores rather than the richer training resources. Therefore, a comparison between $TC_{attn}$ and a model that uses both student essays and the source article is needed.

Acknowledgments

We would like to show our appreciation to every member of the RTA group for sharing their pearls of wisdom with us. We are also immensely grateful to all members of the PETAL group and reviewers for their comments on an earlier version of the paper.

The research reported here was supported, in whole or in part, by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A160245 to the University of Pittsburgh. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

References

Dimitrios Alikaniotis, Helen Yannakoudakis, and Marek Rei. 2016. Automatic text scoring using neural networks. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), volume 1, pages 715–725.

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. Journal of machine Learning research, 3(Jan):993–1022.

Chunshui Cao, Xianming Liu, Yi Yang, Yinan Yu, Jiang Wang, Zilei Wang, Yongzhen Huang, Liang Wang, Chang Huang, Wei Xu, et al. 2015. Look and think twice: Capturing top-down visual attention with feedback convolutional neural networks. In Proceedings of the IEEE International Conference on Computer Vision, pages 2956–2964.

Richard Correnti, Lindsay Clare Matsumura, Laura Hamilton, and Elaine Wang. 2013. Assessing students’ skills at writing analytically in response to texts. The Elementary School Journal, 114(2):142–177.

Scott A Crossley and Danielle S McNamara. 2016. Adaptive educational technologies for literacy instruction. Routledge.
Fei Dong, Yue Zhang, and Jie Yang. 2017. Attention-based recurrent convolutional neural network for automatic essay scoring. In Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017), pages 153–162.

Mohammad K Ebrahimpour, Jiayun Li, Yen-Yun Yu, Jackson Reesee, Azadeh Moghtaderi, Ming-Hsuan Yang, and David C Noelle. 2019. Ventral-dorsal neural networks: Object detection via selective attention. In 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), pages 986–994. IEEE.

Youmna Farag, Helen Yannakoudakis, and Ted Briscoe. 2018. Neural automated essay scoring and coherence modeling for adversarially crafted input. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), pages 263–271.

Corina Florescu and Cornelia Caragea. 2017. Position-rank: An unsupervised approach to keyphrase extraction from scholarly documents. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1105–1115.

Corina Florescu and Wei Jin. 2018. Learning feature representations for keyphrase extraction. In Thirty-Second AAAI Conference on Artificial Intelligence.

Peter W Foltz and Mark Rosenstein. 2015. Analysis of a large-scale formative writing assessment system with automated feedback. In Proceedings of the Second (2015) ACM Conference on Learning@ Scale, pages 339–342. ACM.

Cancan Jin, Ben He, Kai Hui, and Le Sun. 2018. Tdnn: a two-stage deep neural network for prompt-independent automated essay scoring. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 1088–1097.

Beata Beigman Klebanov, Nitin Madnani, Jill Burstein, and Swapna Somasundaran. 2014. Content importance models for scoring writing from sources. In Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pages 247–252.

Tao Lei, Regina Barzilay, and Tommi Jaakkola. 2016. Rationalizing neural predictions. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 107–117.

Xia Li, Minping Chen, Jianyun Nie, Zhenxing Liu, Ziheng Feng, and Yingdan Cai. 2018. Coherence-based automated essay scoring using self-attention. In Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data, pages 386–397. Springer.

Nitin Madnani, Jill Burstein, Norbert Elliot, Beata Beigman Klebanov, Diane Napolitano, Slava Andreyev, and Maxwell Schwartz. 2018. Writing mentor: Self-regulated writing feedback for struggling writers. In Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations, pages 113–117.

Debanjan Mahata, John Kuriakose, Rajiv Ratn Shah, and Roger Zimmermann. 2018. Key2vec: Automatic ranked keyphrase extraction from scientific articles using phrase embeddings. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers), pages 634–639.

Rui Meng, Sanqiang Zhao, Shuguang Han, Daqing He, Peter Brusilovsky, and Yu Chi. 2017. Deep keyphrase generation. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 582–592.

Zahra Rahimi and Diane Litman. 2016. Automatically extracting topical components for a response-to-text writing assessment. In Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications, pages 277–282.

Rod D Roscoe, Laura K Allen, Jennifer L Weston, Scott A Crossley, and Danielle S McNamara. 2014. The writing pal intelligent tutoring system: Usability testing and development. Computers and Composition, 34:39–59.

Antonette Shibani, Simon Knight, and Simon Buckingham Shum. 2019. Contextualizable learning analytics design: A generic model and writing analytics evaluations. In Proceedings of the 9th International Conference on Learning Analytics & Knowledge, pages 210–219. ACM.

Kaveh Taghipour and Hwee Tou Ng. 2016. A neural approach to automated essay scoring. In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pages 1882–1891.

Yi Tay, Minh C Phan, Luu Anh Tuan, and Siu Chuen Hui. 2018. Skipflow: incorporating neural coherence features for end-to-end automatic text scoring. In Thirty-Second AAAI Conference on Artificial Intelligence.

Bronwyn Woods, David Adamson, Shayne Miel, and Elijah Mayfield. 2017. Formative essay feedback using predictive scoring models. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 2071–2080. ACM.

Peng Yuxin, He Xiangteng, and Zhao Junjie. 2018. Object-part attention model for fine-grained image classification. IEEE transactions on image processing: a publication of the IEEE Signal Processing Society, 27(3):1487–1500.
Haoran Zhang and Diane Litman. 2017. Word embedding for response-to-text assessment of evidence. In *Proceedings of ACL 2017, Student Research Workshop*, pages 75–81.

Haoran Zhang and Diane Litman. 2018. Co-attention based neural network for source-dependent essay scoring. In *Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 399–409.

Haoran Zhang, Ahmed Magooda, Diane Litman, Richard Correnti, Elaine Wang, LC Matsmura, Emily Howe, and Rafael Quintana. 2019. erevise: Using natural language processing to provide formative feedback on text evidence usage in student writing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 9619–9625.
A Appendices

A.1 Topic Words Results
Table 9 shows all topic words for the $RTA_{MVP}$ from $TC_{manual}$. Table 10 shows all topic words for the $RTA_{MVP}$ from $TC_{lda}$. Table 11 shows all topic words for the $RTA_{MVP}$ from $TC_{pr}$. Table 12 shows all topic words for the $RTA_{MVP}$ from $TC_{attn}$.

A.2 Specific Example Phrases Results
Table 13 shows all specific example phrases for the $RTA_{MVP}$ from $TC_{manual}$. Table 14 shows all specific example phrases for the $RTA_{MVP}$ from $TC_{lda}$. Table 15 shows all specific example phrases for the $RTA_{MVP}$ from $TC_{pr}$. Table 16 shows all specific example phrases for the $RTA_{MVP}$ from $TC_{attn}$.
| Topic 1  | Topic 2 | Topic 3  | Topic 4  |
|---------|---------|----------|----------|
| care    | bed     | farmer   | school   |
| health  | net     | fertilizer | supplies |
| hospital | malaria | irrigation | fee      |
| treatment | infect  | dying    | student  |
| doctor  | bednet  | crop     | midday   |
| electricity | mosquito | seed     | meal     |
| disease  | bug     | water    | lunch    |
| water    | sleeping | harvest  | supply   |
| sick     | die     | hungry   | book     |
| medicine | cheap   | feed     | paper    |
| generator | infect  | food     | pencil   |
| no       | biting  |          | energy   |
| die      |         |          | free     |
| kid      |         |          | children |
| bed      |         |          | kid      |
| patient  |         |          | go       |
| clinical |         |          | attend   |

Table 9: Topic words of $TC_{\text{manual}}$. 
| Topic 1 | Topic 2 | Topic 3 | Topic 4 | Topic 5 | Topic 6 | Topic 7 | Topic 8 | Topic 9 |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| help    | kenya   | poverty | food    | money   | school  | people  | hospital| years   |
| poor    | like    | think   | fertilizer | need   | kids    | sauri   | medicine| africa  |
| world   | better  | author  | crops   | nets    | supplies | malaria | hospitals| project |
| good    | know    | lifetime | water   | thing   | children | sick    | water   | villages |
| things  | life    | article | farmers | afford  | schools  | 2008    | free    | sauri   |
| time    | help    | possible | needed  | donate  | lunch   | disease | electricity| village |
| work    | think   | convinced | grow    | right   | education | 2004    | diseases| helped  |
| hard    | sauri   | fight   | dying   | dollar  | afford   | nets    | medicines| change  |
| going   | live    | poverty | problem | treatment | energy | mosquitoes | doctors | lives   |
| alot    | clothes | said    | family  | survive  | learn   | getting | 2008    | goals   |
| reason  | states  | achievable | families | needs  | students | says    | gave    | improved|
| happen  | place   | time    | stop    | stuff   | went    | years   | doctor  | 2015    |
| helping | health  | convince | lack    | person  | adults  | progress | examples | help    |
| goal    | important | believe | hunger  | cause   | fees    | died    | 2004    | changed |
| believe | feel    | hannah  | tools   | patients | parents | text    | shape   | year    |
| problems| happy   | shows   | seeds   | provide | 2004    | away    | cure    | changes |
| countries| tell    | reasons | plants  | cost    | lunches | mosquitos | running | started |
| difference| care    | convincing | fertilizers | beds    | books   | prevent | treat   | great   |
| places  | shoes   | fighting | farming | means   | home    | treated | support | millennium|
| change  | story   | wrote   | able    | dont    | wanted  | dieing  | common  | progress |
| little  | america | story   | solved  | dollars | chores  | said    | beds    | came    |
| improve | ways    | agree   | supply  | medical | meal    | come    | patients| girl    |
| country | wants   | saying  | irrigation | jobs   | wood    | night   | said    | 2025    |
| achieve | makes   | opinion | wont    | everyday | materials | bite   | generator | place   |
| hope    | clothing | winning | afford  | gone    | learning | death   | clean   | program |
| helps   | community | Sachs   | hungry  | doctors  | able    | sleep   | electricity| tells   |
| everybody| economy | progress | plant   | lots    | supplies | impoverished | giving | small   |
| start   | history  | conclusion | look    | sickness | meals   | living   | drink   | millenium|
| easy    | paragraph | says    | farms   | live    | paper   | amazing | cures   | read    |
| making  | thats    | future  | feed    | fact    | attendance | easily  | evidence | happened|

Table 10: Topic words of $T_{lda}$. 
| Topic 1 | Topic 2 | Topic 3 | Topic 4 | Topic 5 | Topic 6 | Topic 7 | Topic 8 | Topic 9 | Topic 10 | Topic 11 | Topic 12 | Topic 13 | Topic 14 | Topic 15 | Topic 16 | Topic 17 | Topic 18 | Topic 19 |
|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| irrigation | fertilizer | road | disease | medicine | adults | light | development | villages | project | joy | people | kids | middle | school | fees | village | millennium | backs | plenty | access | care | medicine | schools | today | supply | areas | kind | family |
| farmers | bright | future | malaria | disease | lifetime | villages | project | goals | plan | economics | quality | supporters | dancing | help | health | advice | items | targets | death | night | costs | die | knowledge | food | parents |
| crops | plant | seed | mosquitoes | sick | school | kids | village | millennium | backs | plenty | access | care | medicine | schools | today | supply | areas | kind | family | dancing | help | health | advice | items | targets | death | night | costs | die | knowledge | food | parents |
| Hannah | car | sauri | disease | disease | village | millennium | backs | plenty | access | care | medicine | schools | today | supply | areas | kind | family | dancing | help | health | advice | items | targets | death | night | costs | die | knowledge | food | parents |
| outcome | market | year | time | place | years | life | communities | leaders | glimpse | africa | chemicals | solutions | millions | dancing | help | health | advice | items | targets | death | night | costs | die | knowledge | food | parents |

Table 11: Topic words of $T_{C_{pr}}$. 
| Topic 1   | Topic 2   | Topic 3   | Topic 4   | Topic 5   | Topic 6   | Topic 7   | Topic 8   | Topic 9   | Topic 10  | Topic 11  | Topic 12  | Topic 13  | Topic 14  | Topic 15  | Topic 16  |
|----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| poverty  | way       | years     | lunch     | goals     | electricity| supplies  | afford    | many      | free      | school    | hospital  | bed       | project   | supply    | fertilizer|
| fight    | would     | four      | serves    | problems  | water      | food      | lifetime  | people    | medicine  | plants    | 2004      | nets      | world     | maintain  | seeds     |
| winning  | rate      | villages  | parents   | day       | generator  | net       | could    | kenya     | crops     | fees      | disease   | used      | millennium| diseases  | address   |
| attendance| help      | progress  | passed    | cloth     | running    | rooms     | achievable| sauri     | charge    | students  | yala      | every     | village   | hunger    | irrigation|
|           | kids      | last      | three     | also      | packed     | together  | pencils   | farmers   | africa    | medicines | site      | work      | adults    | tools     |
|           | enough    | occurred  | books     | energy    | needed     | patients  | malaria   | yet       | medicines|          |           |           |          |          |          |
|           | better    | year      | 2015      | connected | 5         | future    | sachs     |           |           |          |           |           |          |          |          |
|           | go        | changes   | knowledge |           |           | keep      | worked    | though    |           |          |          |           |          |          |          |
|           | get       | outcome   | learn     |           |           | poor      | care      | feed      |           |          |          |           |          |          |          |
|           | place     | today     | one       |           |           | five      | family    | two       |           |          |          |           |          |          |          |
|           | solutions | first     |           |           |           | like      | hard      | health    |           |          |          |           |          |          |          |
|           | really    | along     |           |           |           | come      | good      | set       |           |          |          |           |          |          |          |
|           | targets   |          |           |           |           | little    | doctor    | crisis    |           |          |          |           |          |          |          |
|           | see       | treatment |           |           |           | minimal   | either     | areas     |           |          |          |           |          |          |          |
|           | die       | minimal   |           |           |           | whole     | items     |           |           |          |          |           |          |          |          |
|           | hungry    | harvest   |           |           |           | almost   | save      |           |           |          |          |           |          |          |          |
|           | dancing   | showed    |           |           |           | easy      |           |           |           |          |          |           |          |          |          |
|           | walked    | cheap     |           |           |           | met       |           |           |           |          |          |           |          |          |          |
|           | bare      | ever      |           |           |           |           |           |           |           |          |          |           |          |          |          |
|           | feet      | around    |           |           |           |           |           |           |           |          |          |           |          |          |          |
|           | hannah    | mosquitoes|           |           |           |           |           |           |           |          |          |           |          |          |          |
|           | impoverished| encouraging|           |           |           |           |           |           |           |          |          |           |          |          |          |
|           | probably  |          |           |           |           |           |           |           |           |          |          |           |          |          |          |

Table 12: Topic words of $T_{attr}$. 
| Category 1          | Category 2               | Category 3                                    | Category 4                                           |
|---------------------|--------------------------|-----------------------------------------------|-----------------------------------------------------|
| unpaved roads       | united nations intervention | yala sub district hospital                    | malaria common disease preventable treatable         |
| tattered clothing   | safer healthier better life | three kids bed two adults rooms packed patients | mosquitoes carry malaria infect people biting        |
| bare feet           | out poverty stabilize economy quality life communities | not medicine treatment could afford            | kids die malaria adults sick 20,000 day             |
| less than 1 dollar day | africa kenya sauri      | no doctor only clinical officer running hospital | bed nets mosquitoes away people save millions lives |
|                     | goals met 2015 2025      | no running water electricity                  | bed nets cost 5 dollar                               |
|                     | 80 villages across sub-sahara africa | sad people dying near death preventable      | cheap medicines treat malaria                        |
|                     |                          |                                               |                                                     |
| Category 5          | Category 6               | Category 7                                   | Category 8                                          |
| crops dying         | kids not attend go school| progress just four years                      | progress encouraging supporters                      |
| not afford fertilizer irrigation | not afford school fees   | yala sub district hospital has medicine       | solutions problems keep people impoverished         |
| outcome poor crops  | kids help chores fetching water wood | medicine free charge                        | change poverty stricken areas good                  |
| lack fertilizer water | schools minimal supplies books paper pencils | water connected hospital                     | poverty history not easy task hard                  |
| enough food crops harvest feed whole family hungry sick | concentrate not energy | hospital generator electricity                 | winning against poverty possible achievable lifetime |
|                     | no midday meal lunch     | bed nets used every sleeping site             |                                                     |
|                     |                          | hunger crisis addressed fertilizer seeds      |                                                     |
|                     |                          | tools needed maintain food supply             |                                                     |
|                     |                          | kids go school now                            |                                                     |
|                     |                          | no school fees                                |                                                     |
|                     |                          | now serves lunch students                     |                                                     |
|                     |                          | school attendance rate way up                 |                                                     |

Table 13: Specific example phrases of $T_{textual}$. 
| Category 1 | Category 2 | Category 3 | Category 4 | Category 5 |
|------------|------------|------------|------------|------------|
| work hard  | full-time  | author convinces winning fight poverty achievable lifetime | children adults | easy task |
| better place| united states | author convinces winning fight poverty achievable lifetime | disease called malaria | fixed dollar |
| better health| life communities | author convinces winning fight poverty achievable lifetime | come night | thing history |
| brighter future | like books paper parch | winning fight poverty achievable lifetime | malaria mosquitoes | stuff food |
| things like | learn life keep a | winning fight poverty achievable lifetime | easily adults sick | earn money |
| things need | important kids | article states | solutions problems people impoverished | |
| fighting poverty | things important | winning fight poverty achievable | mosquitoes away | |
| hard work | wants know | winning fight poverty achievable | infect people biting | |
| age or author | working hard | article states | away sleeping | |
| working hard | bored | author provided | | |
| author convince | winning fight poverty | author thinks | | |
| winning fight poverty | convincing | based article | | |
| achievable lifetime | poverty achievable lifetime | | | |
| reading article | things changed | | | |

| attendance rate | amazing progress years | good shape | kids adults | donate money |
| midday meal | text says | good education | 2015-2025 | targeted clothing |
| serves lunch students | text said | went school | hungry sick | targeted clothing |
| midday meals | year girl | areas good | cheap medicines | |
| served lunch | year 2004 | trying help | goals supposed | |
| students wanted lunch | paragraph says | worked hard | | |
| books pencils | progress shows winning fight poverty achievable | second reason | | |
| kids attend school | paragraph states | second example | | |
| schools minimal | progress encouraging supporters millennium villages | girl went | | |
| schools hospitals | | | | |
| school school fees | | | | |
| practical items | | | | |
| kids want attend school parents afford school fees | | | | |
| attendance rate | | | | |
| promote money | | | | |

| Category 11 | Category 12 | Category 13 | Category 14 | Category 15 |
|-------------|-------------|-------------|-------------|-------------|
| clean water | grow crops | millennium village project | stop poverty | running water electricity |
| water wood | needed help | millennium village project | long time | water connected hospital generator electricity |
| fresh water | famers worry | millennium village project helped | world work change | patients afford |
| masks help | | change dramatically | beat poverty | rooms packed patients probably | |
| medicines free charge | crops dying affordable necessary fertilizer irrigation | dramatic changes occurred villages subsaharan africa | ending poverty | | |
| crops dying affordable necessary fertilizer irrigation | | place free | | | |
| fertilizer knowledge | hunger crisis addressed fertilizer seeds tools needed maintain food supply | happened years | | | |
| hunger crisis addressed fertilizer seeds tools needed maintain food supply | | | | | |
| for families | dramatic changes occurred villages | | | | |
| hunger crisis addressed | | | | | |
| family plant seeds | | | | | |
| outcome poor | | | | | |
| farmers wanted | | | | | |

| Category 16 | Category 17 | Category 18 | Category 19 | Category 20 |
|-------------|-------------|-------------|-------------|-------------|
| yala subdistrict hospital medicine free charge common diseases | nets sleeping site earn | plan people poverty | achieve goal | years later |
| yala district | afford nets | stabilizes economy quality life communities | reach goal | took years | |
| poor sustainable | | assure access healthcare help | going school | started 2004 | |
| common area | | people people | story says | | |
| diseases like | | wear deaths | | | |
| common disease africa | | poor crops lack | | | |
| hospital good shape | | homeless people | | | |
| divided hospital | | | | | |

Table 14: Specific example phrases of $T^{C_{lda}}$. 
Table 15: Specific example phrases of $TC_{pr}$. 

| Category 1 |
|------------|
| brighter future hannah |
| millennium villages project |
| unpaved dirt road |
| bar sauri primary school |
| future hannah |
| sauri primary school |
| villages project |
| millennium development goals |
| village leaders |
| dirt road |
| car jump |
| little kids |
| preventable diseases people |
| many kids |
| diseases people |
| kids die |
| school supplies |
| primary school |
| school fees |
| infect people |
