Multilingual Short Text Responses Clustering for Mobile Educational Activities: a Preliminary Exploration

Yuen-Hsien Tseng[1], Lung-Hao Lee[1,2], Yu-Ta Chien[3], Chun-Yen Chang[4], Tsung-Yen Li[4]
[1]Graduate Institute of Library and Information Studies, National Taiwan Normal University
[2]MOST Joint Research Center for AI Technology and All Vista Healthcare, NTU
[3]Institute of Education, National Taiwan Ocean University
[4]Science Education Center, National Taiwan Normal University
{samtseng, changcy, yan}@ntnu.edu.tw
lhlee@ntu.edu.tw, ytchien@ntou.edu.tw

Abstract

Text clustering is a powerful technique to detect topics from document corpora, so as to provide information browsing, analysis, and organization. On the other hand, the Instant Response System (IRS) has been widely used in recent years to enhance student engagement in class and thus improve their learning effectiveness. However, the lack of functions to process short text responses from the IRS prevents the further application of IRS in classes. Therefore, this study aims to propose a proper short text clustering module for the IRS, and demonstrate our implemented techniques through real-world examples, so as to provide experiences and insights for further study. In particular, we have compared three clustering methods and the result shows that theoretically better methods need not lead to better results, as there are various factors that may affect the final performance.

1 Introduction

The development of Natural Language Processing (NLP) has advanced to a level that affects the research landscape of academic domains and has practical applications in various industrial sectors. On the other hand, educational environment has also been improved to impact the world society, such as the emergence of MOOCs (Massive Open Online Courses), and new learning tools or teaching paradigms have also change the way of class interactions, such as the use of Classroom Response Systems (CRS) (Siau et al., 2006). The advance of these two fields has converged to support some of the online or on-site course activities that are previously infeasible, such as real-time understanding of student's responses (Beatty and Gerace, 2009) and mobile language learning (Cardoso, 2010).

Research issues in this direction have gained more and more attention (Hearst, 2015). Examples include the workshops on Innovative Use of NLP for Building Educational Applications (BEA) since 2003[1] and the workshops on Natural Language Processing Techniques for Educational Applications (NLPTEA) since 2014[2], where the former was held in North America mainly for English or western languages, while the latter was held in Asia mainly for Chinese or oriental languages.

NLP for educational applications not only concerns the academic community, but also has great potential in the educational market. Systems for online writing evaluation service (or automated essay scoring) like ETS's Criterion[3] and for plagiarism identification like Turnitin[4] have established their market share. However, these successful services are built upon mature educational activities and deal with relatively long articles or complete sentences for reliable performance. In contrast, processing of short texts (or sub-sentences, non-sentences, or even a few terms) is under-developed for novel educational applications.

[1] https://ekaterinakochmar.wixsite.com/sig-edu
[2] https://www.sigcall.org/
[3] http://www.criterion.com.tw/
[4] http://turnitin.com/zh_tw/

Proceedings of the 5th Workshop on Natural Language Processing Techniques for Educational Applications, pages 157–164
Melbourne, Australia, July 19, 2018.

Electronic classroom response systems (CRS), also called instant response systems (IRS) or clickers, have been tested and used in higher education classrooms since the 1960's (Deal, 2007). According to a CNET report (Gilbert, 2005), schools and universities, most in the United States, bought nearly a million clickers in year 2004 alone, using infrared or radio frequency technology for students' transmitters. This number accumulated to nearly nine million units in under a decade by just two of many companies that make clickers (Hoffman, 2012). Recently, IRS has gained even greater popularity in class interaction (Bartsch and Murphy, 2011; Chen et al., 2013; Han, 2014; Morais et al., 2015) due to the ubiquitous availability of mobile devices for each individual and cloud-based technology for ease of data collection and integration. IRS services in Taiwan like Zuvio (http://www.zuvio.com.tw/) have attracted local university users in a short term because of its easier use than traditional transmitter-required IRS and LMS (Learning Management System) such as Moodle App. For example, over the course of year 2014, Zuvio usage in National Taiwan University (NTU) increased from 61 to 263 instructors, 68 to 384 courses, and 2,037 to 11,172 students (Lee and Shih, 2015).

By broadcasting a question to all students' mobile devices and getting responses instantly, such systems help teachers know the learning status of each student better and also help students maintain their attention during the class due to the instant feedback from the teachers and/or their classmates (Bartsch and Murphy, 2011; Beatty and Gerace, 2009). However, the potential of such IRS may still be under-explored (Chien and Chang, 2015a). In the above NTU case, although the majority (54%) of questions deployed in Zuvio were multiple choices, many instructors also used open-ended questions (20%) and composite questions (21%) to promote deeper engagement and reflection (J. W.-S. Lee and Shih, 2015). Previous studies even indicated that multiple-choice examinations pose an obstacle for higher-level thinking in science classes (Stanger-Hall, 2012) and constructed response (e.g. free text writing) assessments are widely viewed as providing greater insight into student thinking than closed form (e.g. multiple-choice) assessments (Birenbaum and Tatsuoka, 1987).

However, no IRS system has yet provided analysis of these open-ended text responses in real time, to our best knowledge. By applying NLP techniques to the IRS or similar mobile interaction systems where only short text interaction is feasible, more information for the students could be provided and therefore more meaningful engagement and efficient learning could be achieved (Chien and Chang, 2015b).

Based on the above trends and observations, this study aims at developing related NLP techniques applicable to the current and future educational environment. More specifically, this paper focuses on the short text response processing in the situation where some forms of instant response systems (IRS) are used in and after the class.

2 Short Text Response Clustering

As our purpose is to support IRS-related educational activities, an existing IRS would be used for integrating the techniques to be developed so that we can focus on the required new functions without re-inventing the wheel. We choose CloudClassRoom (CCR, http://ccr.tw/) because it is developed by the team of our collaborators (Chien and Chang, 2015a) and because it supports at least 12 languages for international use. This choice would facilitate our testing and evaluation of the developed techniques. However, we keep in mind that the techniques to be developed should be independent of the CCR system, such that they can be ported to another IRS instantly. In fact, CCR is developed in JQuery and PHP language, while the NLP techniques to be developed mainly use Python as our programming language.

Once we have an IRS platform, we can package the required techniques into one of the IRS's module to meet the research purposes. Figure 1 shows a series of processing step packaged into a Semantic Processing Module (SPM), where each rectangular box denotes a processing sub-module and each cylinder denotes a set of language knowledge, corpora, resources, or technical options.

The first-row in the figure mainly deals with refining the terms from the response texts, which heavily depends on the language knowledge and resources. The second-row deals with the sematic processing of the texts, which is basically language independent, except the term expansion step. This pipeline structure is inevitable as there are many options in processing texts for a certain

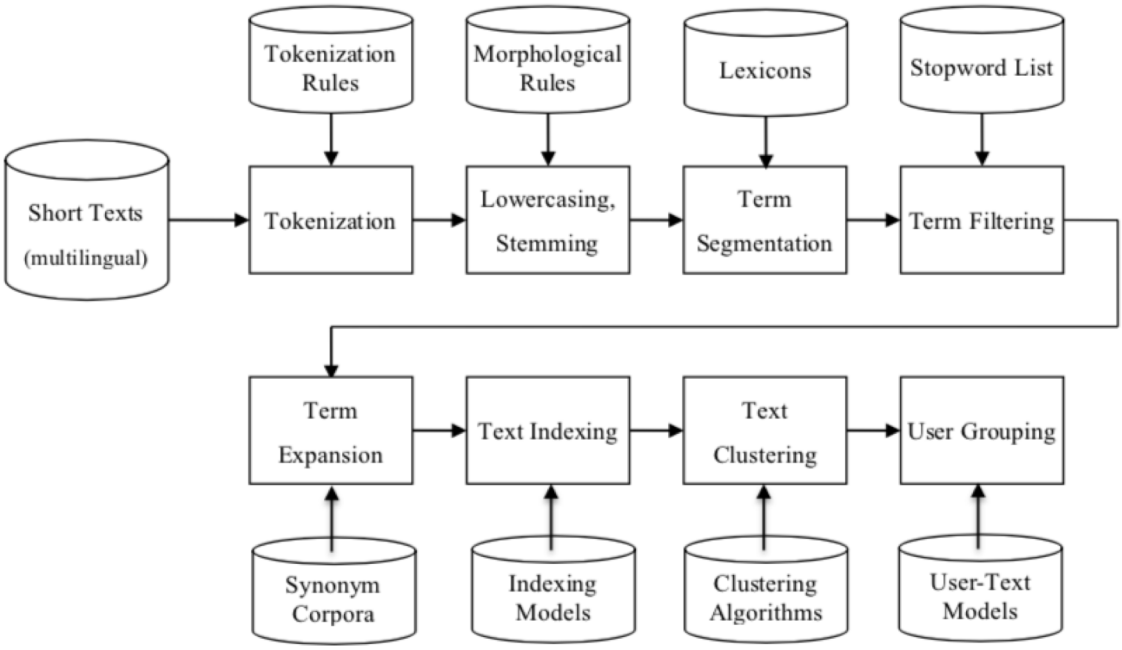

Figure 1: Pipeline steps of the SPM for processing short texts in an IRS system**.**

task in a certain language. At our early stage of development, each step would have options for selection by teachers or by NLP experts to best suit the educational activities in a certain course. At the later stage, we expect that the SPM should finally learn the options without human selection. For example, the tokenization need to transform all different digital numbers into a single numeric symbol for semantic clustering in general cases, but should leave the numbers intact in courses such as mathematics, where exact numbers from students are expected for accurate processing. The case also applies to the morphological step where lowercasing and stemming are applied for English semantic processing in general cases, but the morphological analysis should be turned off when, e.g., English is taught, or the expected answers are exact terms used by the students. This consideration would optimize the SPM for each educational activity, but may require years of fine-tuning when more and more activities are encountered in real-world applications. In fact, the CCR has at least 4780 teachers registered, 11,784 classrooms established, and 23,376 questions asked and 248,633 responses received. It really contains many valuable resources for NLP experiments and applications.

3 Demonstration

To have a concrete idea about the texts submitted by students via CCR, Table 1 shows a set of real-world texts in response to the question asked by a Taiwan university teacher of General Education of Science: "As a marine researcher, if someone presents the photos shown in Figure 2 to you and ask your opinion about the creature, what would you think of and what would you ask?"

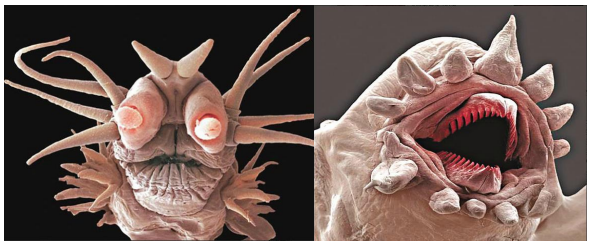

Figure 2: Photographs to trigger questions for students to respond.

As can be seen from Table 1, there are several characteristics in the students' responses: 1) meaningfulness punctuations, e.g., ID 3; 2) multi-lingual: English responses even in a Chinese

Student ID	In this table, there are 29 student responses to the question: "As a marine researcher, if someone presents the photos shown in figure to you and ask your opinion about the creature, what would you think of and what would you ask?"
1	1.發現地點 2. 推論有毒 3. 外星生物 地球沒有
3	1.好吃嗎。2.肉食性。3.牙齒很尖。4.深海魚。5.因為很醜==
4	What's its life cycle? I guess that it's a meat-eater. Maybe it's a parasite. Because it has a fixation structure.
5	他們有毒嗎？
6	他們看得到嗎？他們的食物可能是什麼？
9	他是生活在何種海域?!深度?!環境?!
11	它的體型大小
12	住在很深的海裡吧！眼睛很凸。牙齒很尖應該是食物動物！
15	呵呵呵呵呵
17	問題：在哪個海域發現的呢？特性：吃腐肉，極可能是古老的活化石。原因：場項特別、牙齒尖銳。生物：恐龍？！或其相關生物
18	問題：水深位置大約在何處。推測：疑似刺絲胞動物門，有攻擊力，有尖銳的外型
21	好奇怪
22	它有肛門嗎？
23	對光源有無反應
24	很像大英雄天團的噴火龍
25	我覺得它是人類的祖先。因為他有眼睛 有嘴巴 有牙齒。牠以細菌為主，牠屬於夜行性動物，睡眠時間為 12 小時 也就是半天，是個奇怪的生物！！！我想要 usb
27	海洋生物
28	深海生物，無視覺
29	爸爸
31	牠生活於海洋表層還深層？狩獵能力較強，因為牙齒以犬齒較多可進行撕裂，海洋深層消費者
32	發現的環境包含深度 身體外表特徵，實行生物分類 它的捕食習性
33	眼睛會感光嗎？肉食性動物，牙齒看起來很尖，深海的未知生物，因為看不出來是什麼種類的生物
34	神奇寶貝
35	肉食性的魚，在很深暗海
36	觸角是類似珊瑚的觸角嗎。應該是住在深海裡的雙種生物吧。
37	跟我同學很像
38	身體構造有那些特徵
39	這個生物是不是小小隻的？可能是吃浮游生物的，深海的生物，因為有觸手
43	這種生物有攻擊性嗎？應該住在深海？有照明的能力吧！這會不會是鯊魚和燈籠魚的合體

Table 1: Examples of text responses from students via CCR.

class, e.g., ID 4; 3) nonsense responses, e.g. ID 15, 24, 29, etc.; 4) very short texts, e.g., ID 5, 11, 27, etc.; and 5) non-topical texts, e.g. the last part of ID 25, where the student asks for a prize promised by the teacher who encourages the students to aggressively respond to the question for a USB storage device as a prize.

Characteristic 1 can be removed at the tokenization stage. Characteristic 2 could be translated using simple word-by-word translation (by way of multi-lingual lexicons or multilingual WordNets[5], such as BabelNet[6]), with translation tools such as Goslate[7], or customized machine translation techniques (Chuang and Tseng, 2008; Tseng et al., 2011). Characteristic 4 can be extended by synonym lexicons or multilingual WordNets to enrich the textual information. However, despite we have eHownet[8] resources from the ACLCLP (Association of Computational Linguistics and

[5] https://wordnet.princeton.edu/
[6] https://babelnet.org/
[7] https://pythonhosted.org/goslate/
[8] http://ehownet.iis.sinica.edu.tw/index.php

Chinese Language Processing), there is no guarantee that the synonyms or hypernyms in eHownet is able to cover the terms used in a class like the above. After these preprocessing, Characteristics 2, 3, 4, and 5 require an effective text clustering technique to distinguish them from the normal meaningful responses, such that the teacher could decide what to do for the improper responses. Once they can be isolated in real time, the teacher can, for example, ask the corresponding students to re-submitted their responses, or preset the system to prevent these texts from been submitted by the students.

To have an idea of how well existing clustering techniques can do for these texts, we have tried three approaches:

(1) Hierarchical Agglomerative Clustering (HAC) based on a hybrid way of term indexing, namely lexicon-based segmentation followed by a keyword extraction using the method of (Tseng, 1998, 2002; Tseng et al., 2010b), implemented in a well-debugged tool called CATAR (Tseng, 2010a; Tseng and Tsay, 2013), as shown in Figure 3.

(2) Latent Semantic Analysis (LSA) based on the word segmentation by jieba and a topic modeling tool genism without removal of any stopwords and punctuations, as shown in Figure 4.

(3) Latent Dirichlet Analysis (LDA) by jieba and gensim with stopwords and punctuations being removed, as shown in Figure 5.

From Figure 3 based on HAC, there are 3 multi-documents clusters and 16 singleton clusters. The result is generally reasonable, only a few texts, like ID 23 and 31, could not be grouped together with other similar texts. This is because a rigorous criterion is imposed on the HAC, i.e., complete linkage clustering such that ID 31 did not cluster into Cluster 3, despite it contains the salient term "牙齒" in Cluster 3. Also, the lexicon-based segmentation regards "深海", "海洋", and "海洋深層" as different terms, such that they are totally different features for text clustering. The above two reasons may also apply to the terms and texts, such as "光源" (ID 23), "感光" (ID 33), "暗海" (ID 35), and "夜行性" (ID 25), or "食物" (ID 6) and "肉食性" (ID 3, 33, and 35).

Figure 3: HAC clustering results.

Group ID	Student ID:
#1	3：1 好吃嗎。2 肉食性。3 牙齒很尖。… 17：問題：在哪個海域發現的呢？… 18：問題:水深位置大約在何處。… 31：牠生活於海洋表層還深層 ？… 33：眼睛會感光嗎？… 36：觸角是類似珊瑚的觸角嗎，… 39：這個生物是不是小小隻的？… 43：這種生物有攻擊性嗎？…
#2	5: 他們有毒嗎? 6: 他們看得到嗎? 他們的食物可能是什麼?
#3	11: 他的體型大小 25: 我覺得它是人類的祖先。… 32: 發現的環境包含深度 … 37: 跟我同學好像 38: 身體構造有那些特徵
#4	1: 發現地點 … 4: What's its life cycle? … 9: 他是生活在何種海域?! … 22: 它有肛門嗎? 23: 對光源有無反應
#5	12: 住在很深的海裡吧! 眼睛很凸 … 24: 很像大英雄天團的噴火龍 28: 深海生物，無視覺 35: 肉食性的魚，在很深暗海
N/A	15: 呵呵呵呵呵 21: 好奇怪 27: 海洋生物 29: 爸爸 34: 神奇寶貝

Figure 4: LSA clustering results.

Group ID	Student ID
#1	1: 1. 發現地點 2. 推論有毒 … 5: 他們有毒嗎？ 6: 他們看得到嗎? … 21: 好奇怪 22: 它有肛門嗎？
#2	31: 牠生活於海洋表層還深層？ 39: 這個生物是不是小小隻的？ 37: 跟我同學好像 27: 海洋生物
#3	12: 住在很深的海裡吧！… 33: 眼睛會感光嗎？肉食性動物… 43: 這種生物有攻擊性嗎？… 36: 觸角是類似珊瑚的觸角嗎，… 9: 他是生活在何種海域?! ... 15: 呵呵呵呵呵
#4	18: 問題:水深位置大約在何處。… 28: 深海生物，無視覺 3: 1 好吃嗎。2 肉食性。3 牙齒很尖。… 23: 對光源有無反應 32: 發現的環境包含深度 … 38: 身體構造有那些特徵 4: What's its life cycle? … 24: 很像大英雄天團的噴火龍 29: 爸爸
#5	17: 問題:在哪個海域發現的呢? … 35: 肉食性的魚，在很深暗海 11: 他的體型大小 25: 我覺得它是人類的祖先。… 34: 神奇寶貝

Figure 4: LDA clustering results.

To improve the performance such that the texts containing these semantically related terms being clustered together, it seems that LSA or LDA are better solutions as past studies have shown the possibility (Blei et al., 2003; Deerwester et al. 1990). Based on the HAC result, there are 3-5 clusters in this case. So we cluster the responses using 5 topics with LSA and LDA. Actually, this number: about 5 clusters for each set of responses, is a proper choice for science education based on the feedback of our co-investigator. However, Figure 4 and 5 shows that LSA and LDA alone cannot solve this short-text clustering problem better. They can sometimes lead to worse results. In addition to the shortage of textual information (short texts), there are other factors that influence the performance, such as feature extraction (whether to use 1-grams as features in Chinese short texts or not, such as "海", "光"), term expansion (whether to incorporate the term-level similarity, such as those between "感光" and "夜行性", or "食物" and "肉食性", into text clustering). Furthermore, these decisions may depend on the characteristics of the questions asked or classes taught. Therefore, we propose the pipeline SPM in Figure 1 to deal with this problem, so that in each step we could choose proper options for better performance.

To incorporate more semantic information into the SPM, we plan to use language resources such as eHownet, WordNet, and BabelNet for Chinese, English, and multilingual synonym expansion, respectively. Our future study would also use tools like word2vec (Mikolov et al., 2013) and concept map miner (Tseng et al., 2010; Tseng et al., 2012) to extract paradigmatically and/or topically similar terms for term expansion (Tseng et al., 2010). In addition to term expansion, utilization of contextual information of the short texts can be enhanced by machine translation (Tang et al., 2012). Direct clustering based on the continuous distributed representations of words, sentences, or paragraphs (Chinea-Rios et al., 2015; Mikolov et al., 2013) may also be worth of exploring. As a tradition in NLP research, further study will try all the promising combinations of the mentioned techniques to see which combinations perform best in which conditions.

As to the clustering performance evaluation, there are intrinsic and extrinsic measures, where the former measures the clustering quality directly and the latter measures the quality indirectly by applying the clustering result to other task and see if a good result can be obtained from the task. For intrinsic evaluation, measures like perplexity, Rand index, and Silhouette index have been used and we have implemented the latter two measures (Rand and Silhouette) in CATAR to help determine the number of clusters (Tseng, Lin, & Lin, 2007; Tseng & Tsay, 2013). For extrinsic evaluation, which is more suitable for the IRS applications, it depends on how the teacher would like the clustering results. Therefore, our strategies would implement different clustering techniques and intrinsic evaluation measures to suggest various cluster results for the teachers to choose a proper one. Before that, we had assisted the teachers to quickly understand a clustering result by providing some intrinsic evaluation result, i.e., the cluster descriptors as shown in Fig-

ure 3. In this way, we help the teachers to explore the students' responses in a period of time short enough during their lecturing activities using the IRS.

4 Conclusions

This paper describes our preliminary study of short text response clustering for mobile educational activities. We illustrate the characteristics of short text responses from the IRS, propose the SPM module for processing short texts, and demonstrate our implemented techniques via the CCR system. We also compare three clustering methods, and the results showed that theoretically better methods need not lead to better results, as there are various factors that may affect the final performance.

In real-case applications, the SPM module based on the LSA technique has been used online for two years, serving thousands of teachers. Informal evaluation from the responses of teachers, including those in Taiwan and Thailand, has shown that the proposed short-text clustering is applicable to their educational activities.

Acknowledgments

This study was partially supported by the Ministry of Science and Technology (MOST), Taiwan, R.O.C., under the grant: MOST 105-2221-E-003-020-MY2, MOST 106-2221-E-003-030-MY2, and MOST 107-2634-F-002-019-.

References

R. A. Bartsch, and M. Murphy. 2011. Examining the effects of an electronic classroom response system on student engagement and performance. *Journal of Educational Computing Research*, 44(1): 25-33. https://doi.org/10.2190/EC.44.1.b

I. Beatty, and W. Gerace. 2009. Technology-enhanced formative assessment: a research-based pedagogy for teaching science with classroom response technology. *Journal of Science Education and Technology*, 18(2): 146-162. https://doi.org/10.1007/s10956-008-9140-4

M. Birenbaum, and K. K. Tatsuoka. 1987. Open-ended versus multiple-choice response formats—it does make a difference for diagnostic purposes. *Applied Psychological Measurement*, 11(4): 385-395. https://doi.org/10.1177/014662168701100404

D. M. Blei, A. Y. Ng, and M. I. Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.

W. Cardoso. 2010. Clickers in foreign language teaching: a case study. *Contact: Teachers of English as a Second Language of Ontario*, 36(2): 36-55. http://spectrum.library.concordia.ca/36087/

T.-L. Chen, Y.-F. Lin, Y.-L. Liu, H.-P. Yueh, H.-J. Sheen, and W.-J. Lin. 2013. Integrating Instant Response System (IRS) as an in-class assessment tool into undergraduate chemistry learning experience: student perceptions and performance. In M.-H. Chiu, H.-L. Tuan, H.-K. Wu, J.-W. Lin, and C.-C. Chou (Eds.), *Chemistry Education and Sustainability in the Global Age* (pp. 267-275): Springer Netherlands.

Y.-T. Chien, and C.-Y. Chang. 2015a. Providing students with an alternative way to interact with the teacher in the silent classroom: Teaching with the CloudClassRoom technology. In *Proceedings of the Inaugural Asian Conference on Education & International Development*.

Y.-T. Chien, and C.-Y. Chang. 2015b. Supporting socio-scientific argumentation in the classroom through automatic group formation based on students' real-time responses. In M. S. Khine (Ed.), *Science education in East Asia: Pedagogical innovations and research-informed practices* (pp. 549-563): Springer International Publishing.

M. Chinea-Rios, G. Sanchis-Trilles, and F. Casacuberta. 2015. Sentence clustering using continuous vector space representation. In R. Paredes, J. S. Cardoso, and X. M. Pardo (Eds.), *Pattern Recognition and Image Analysis* (Vol. 9117, pp. 432-440): Springer International Publishing.

Z.-J. Chuang, and Y.-H. Tseng. 2008. NTCIR-7 experiments in patent translation based on open source statistical machine translation tools. In *Proceedings of the 7th NTCIR Workshop on Evaluation of Information Access Technologies: Information Retrieval, Question Answering, and Cross-Lingual Information Access*.

A. Deal. 2007. A teaching with technology white paper: classroom response systems. Retrieved from https://www.cmu.edu/teaching/technology/whitepapers/ClassroomResponse_Nov07.pdf

S. Deerwester, S Dumais, G Furnas, T. K. Landauer, and R. Harshman. 1990. Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, *41* (6): 391–407.

A. Gilbert. 2005. New for back-to-school: 'Clickers'. Retrieved from http://www.cnet.com/news/new-for-back-to-school-clickers/

J. H. Han. 2014. Closing the missing links and opening the relationships among the factors: a literature review on the use of clicker technology

using the 3P model. *Journal of Educational Technology & Society, 17*(4): 150-168.

M. A. Hearst. 2015. Can natural language processing become natural language coaching? In *Proceedings of the 53rd Annual Meeting of the Association of Computational Linguistics*.

J. Hoffman. 2012. Speak up? Raise your hand? That may no longer be necessary. Retrieved from http://www.nytimes.com/2012/03/31/us/clickers-offer-instant-interactions-in-more-venues.html?_r=0

J. W.-S. Lee, and M.-l. Shih. 2015. Teaching practices for the student response system at National Taiwan University. *International Journal of Automation and Smart Technology*, 5(3): 145-150. https://doi.org/10.5875/ausmt.v5i3.862

T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, amd J. Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of the Advances in Neural Information Processing Systems*.

A. Morais, J. I. Barragués, and J. Guisasola. 2015. Using a classroom response system for promoting interaction to teaching mathematics to large groups of undergraduate students. *Journal of Computers in Mathematics and Science Teaching, 34*(3): 249-271.

K. Siau, S. Hong, amd F. F. H. Nah. 2006. Use of a classroom response system to enhance classroom interactivity. *IEEE Transactions on Education*, 49(3): 398-403. https://doi.org/10.1109/TE.2006.879802

K. F. Stanger-Hall. 2012. Multiple-choice exams: an obstacle for higher-level thinking in introductory science classes. *CBE Life Sciences Education, 11*(3): 294-306. https://doi.org/10.1187/cbe.11-11-0100

J. L. Tang, X. F. Wang, H. J. Gao, X. Hu, and H. Liu. 2012. Enriching short text representation in microblog for clustering. *Frontiers of Computer Science*, 6(1): 88-101. https://doi.org/10.1007/s11704-011-1167-7

Y.-H. Tseng. 1998. Multilingual keyword extraction for term suggestion. In *Proceedings of the 21st International ACM SIGIR Conference on Research and Development in Information Retrieval.*

Y.-H. Tseng. 2002. Automatic thesaurus generation for Chinese documents. *Journal of the American Society for Information Science and Technology, 53*(13): 1130-1138. https://doi.org/10.1002/asi.10146

Y.-H. Tseng. 2010a. Content Analysis Toolkit for Academic Research (CATAR). Retrieved from http://web.ntnu.edu.tw/~samtseng/CATAR/

Y.-H. Tseng. 2010b. Generic title labeling for clustered documents. *Expert Systems With Applications, 37*(3): 2247-2254. https://doi.org/10.1016/j.eswa.2009.07.048

Y.-H. Tseng, C.-Y. Chang, S.-N. R. Chang, and C.-J. Rundgren. 2010. Mining concept maps from news stories for measuring civic scientific literacy in media. *Computers & Education*, 55(1): 165-177. https://doi.org/10.1016/j.compedu.2010.01.002

Y.-H. Tseng, Z.-P. Ho, K.-S. Yang, and C.-C. Chen. 2012. Mining term networks from text collections for crime investigation. *Expert Systems With Applications*, 39(11): 10082-10090. https://doi.org/10.1016/j.eswa.2012.02.052

Y.-H. Tseng, C.-J. Lin, and Y.-I. Lin. 2007. Text mining techniques for patent analysis. *Information Processing and Management*, 43(5): 1216-1247. https://doi.org/10.1016/j.ipm.2006.11.011

Y.-H. Tseng, C.-L. Liu, C.-C. Tsai, J.-P. Wang, Y.-H. Chuang, and J Jeng. 2011. *Statistical approaches to patent translation for patentMT- experiments with various settings of training data*. In *Proceedings of the 9th NTCIR Workshop on Evaluation of Information Access Technologies: Information Retrieval, Question Answering, and Cross-Lingual Information Access*.

Y.-H. Tseng, and M.-Y. Tsay. 2013. Journal clustering of Library and Information Science for subfield delineation using the bibliometric analysis toolkit: CATAR. *Scientometrics, 95*(2): 503-528.