Rodrigo, M. M. T. et al. (Eds.) (2021). Proceedings of the 29th International Conference on Computers in Education. Asia-Pacific Society for Computers in Education

Examining the Effects of Automatic Speech Recognition Technology on Learners' Lexical Diversity

Michael Yi-Chao JIANG*, Morris Siu-Yung JONG & Wilfred Wing-Fat LAU
Department of Curriculum and Instruction & Centre of Learning Sciences and Technologies, The Chinese University of Hong Kong, Hong Kong S. A. R., China
*mjiang@cuhk.edu.hk

Abstract: A total of 160 undergraduates participated in the 14-week quasi-experiment. The experimental group and the control group were both taught with a flipped approach, but the students in the experimental group were required to conduct an additional automatic speech recognition-based pre-class task. The vocd-D value and MTLD were adopted as metrics of students' lexical diversity. A two-way between- and within-subjects repeated measures design was conducted to examine the effects of the group factor, the time factor and the group × time interaction effects. The results showed that the students in the experimental group scored statistically better than their counterparts in the control group on both the vocd-D value and MTLD. However, no significant difference was witnessed over time and there was no significant group × time interaction effect in either group.

Keywords: Automatic speech recognition, flipped classroom, lexical diversity, English as a foreign language

1. Introduction

In flipped learning, students are supposed to prepare themselves with the pre-class content actively and achieve a proper level of preparedness. However, it was revealed that educational technology was not fully harnessed in flipped classrooms (Jiang et al., 2020). Advanced technologies such as automatic speech recognition (ASR) are not commonly considered and utilized. Because of the limited use of technology in some flipped classrooms, students may become as demotivated as do they in traditional classrooms. In EFL research, lexical diversity is considered as a basic descriptor of learners' oral English proficiency. Although many extant studies have reported positive evidence of the effectiveness of the FCA in EFL learning, few studies have adopted domain-specific indicators (e.g., specific measures of students' oral fluency or accuracy) to examine the effectiveness of the technology-enhanced flipped classroom approach (FCA) (Jiang et al., 2021). Therefore, the present study aims to examine the effects of ASR technology on EFL learners' lexical diversity in a flipped setting. Studies with more refined indicators of learners' linguistic performance can contribute to a deeper understanding of the FCA for language learning and diversify the instructional design of the FCA. Accordingly, we formulated two research questions: 1) Do students in a flipped EFL classroom outperform their counterparts in a traditional EFL classrooms in terms of lexical diversity? 2) How does the flipped classroom approach improve the EFL students' lexical diversity over time?

2. Methods

A total of 160 first-year students from a four-year university were enrolled in the main study. Before the course began, the four classes had been randomly assigned into an experimental group (EG) and a control group (CG). The EG students were assigned a mediating ASR-based oral task in addition to the self-learning resources on Unipus for pre-class preparation. In contrast, the CG students were only given the materials on Unipus before class. All the students were randomly assigned to workgroups of

two to four for in-class activities within each class. For each unit, the students in both groups undertook a communicative task in class in English and recorded the whole activity using their mobile phone in an auditory fashion. Within each workgroup, the students orally expressed their opinions or experiences regarding the unit topic. The recordings of Units 2, 4, 6 and 8 were used for data analysis, but the students did not know which unit would be analyzed. After data screening, the present study conducted data analyses based on the data that were transcribed and coded from a total of 128 participants (68 EG students and 60 CG students). The in-class recordings were transcribed into plain text and coded and annotated using ELAN (https://tla.mpi.nl/tools/tla-tools/elan). The vocd-D value and MTLD were computed with TextInspector (Figure 1), a professional website for studying linguistic features of English spoken language. The transcribed recordings of in-class peer interaction were coded into frequencies and relative frequencies (against AS-unit) to form study-generated quantitative data. A mixed within- and between-subjects design (pre-intervention English proficiency controlled for as a covariate) was conducted. The independent variables were the group factor and the time factor. The dependent variables were the metrics of lexical diversity coded from students' in-class task performance.

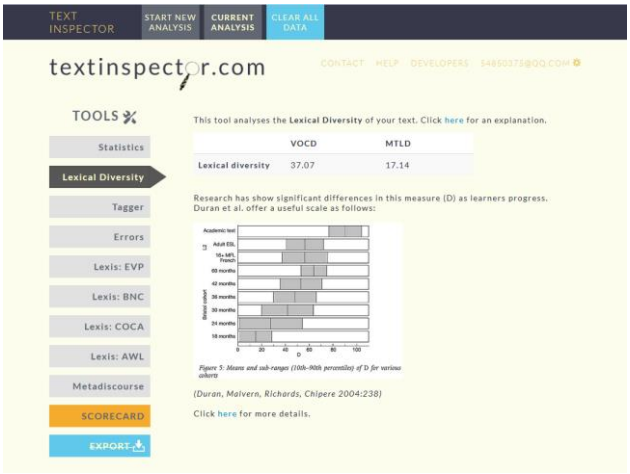

Figure 1. Screenshot of TextInspector for Calculating vocd-D and MTLD.

3. Results and Discussion

The means of vocd-D value over time was 41.716 for the EG students and 36.483 for the CG students. The means of MTLD over time was 32.807 for the EG students and 28.060 for the CG students. In terms of the timepoints, the means of the vocd-D values across the two groups was 35.552 for Time 1, 36.009 for Time 2, 39.782 for Time 3, and 44.401 for Time 4, showing a clear upward tendency (Figure 2). The means of MTLD across the two groups was 28.700 for Time 1, 29.463 for Time 2, 28.497 for Time 3 and 31.958 for Time 4, indicating an upward yet fluctuating pattern. The test of between-subjects effects revealed that the main effect of the group factor on the average score of vocd-D values across time was significantly different ($F_{(1,\ 125)} = 3.945$, $p = 0.049 < 0.05$) and the effect size was small-to-medium ($\eta^2_p = 0.031 > 0.01$). Likewise, the main effect of the group factor on the average score of MTLD across time was also significantly different ($F_{(1,\ 125)} = 6.244$, $p = 0.014 < 0.05$) also with a small to medium effect size ($\eta^2_p = 0.048 > 0.01$). Therefore, the EG students performed significantly better than their counterparts in the CG in terms of lexical diversity.

The tests of within-subjects effects showed that the main effect of time was not statistically significant on the average scores on vocd-D values ($F_{(2.475,\ 309.351)} = 0.483$, $p = 0.658 > 0.05$) or on MTLD ($F_{(2.481,\ 302.224)} = 0.049$, $p = 0.971 > 0.05$), sphericity not assumed. Additionally, the 'group × time' interaction effects on vocd-D value ($F_{(2.475,\ 309.351)} = 2.720$, $p = 0.055 > 0.05$) or on MTLD ($F_{(2.481,\ 302.224)} = 0.827$, $p = 0.458 > 0.05$) were not statistically significant, sphericity not assumed. To sum up, the time factor did not lead to any statistically significant effects on lexical diversity, and there was no group × time interaction effect on neither of them.

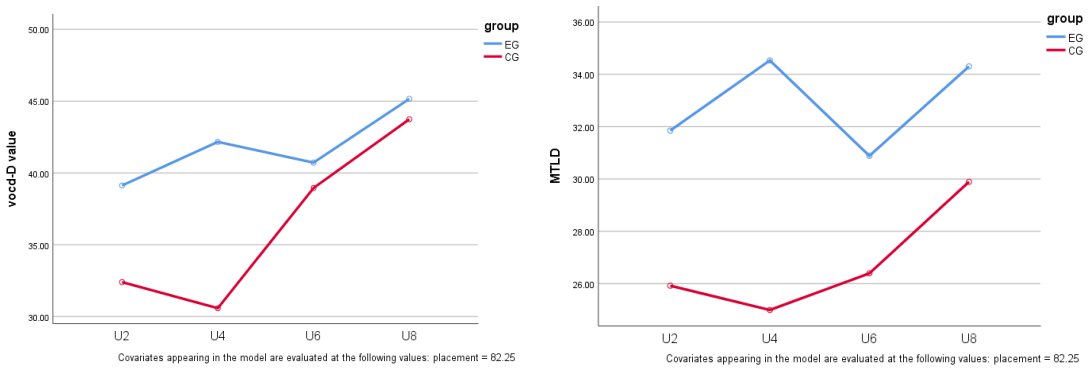

Figure 2. Profile Plots of vocd-D Value and MTLD.

The significant gains on lexical diversity indicated that ASR technology had positive effects on EG students' in-class linguistic performance in terms of lexical diversity in oral English. The EG students were required to use the ASR technology-based iFlyRec to practice their English speaking skills for their pre-class self-learning. They were encouraged to repeatedly practice answering the lead-in questions attached to the reading sections. The students needed to master the new vocabulary and have a good knowledge of the text contents to answer the lead-in questions properly, which to some extent, reinforced their mastery of and facilitated the use of the new words. In contrast, their counterparts in the CG did not have any required practice of using the new vocabulary. According to the responses in the post-intervention interview, it was found that most of the students self-studied the new vocabulary mostly by memorizing by rote, focused only on building meaning connections between Chinese and English, spelling and pronunciation. They had little awareness in learning the usage of the new words. Therefore, when performing the communicative tasks in class, the EG students scored statistically more on the lexical diversity dimension. Students' lexical diversity was not enhanced significantly over time. These resultant contrasting findings align with most of the complexity, accuracy and fluency (CAF) studies conducted among non-native speakers (e.g., Skehan, 2009). It was found in those studies that the correlation between lexical diversity and syntactic complexity was shown to be negative, indicating that for non-native speakers, 'more varied lexis seems to cause problems for non-native speakers and provokes more errors while not driving forward lexical diversity' (Skehan, 2009, p. 116).

4. Conclusion

The results revealed that the EG students scored statistically higher on both vocd-D and MTLD than the CG students, indicating that the EG students produced more complex utterances on the lexical level than their counterparts in the CG. However, regardless of their group membership, students' lexical diversity did not improve significantly over time. In other words, the ASR technology significantly improved the EG students' lexical diversity, but over time, the lexical diversity of both the EG students and the CG students did not improve significantly.

References

Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., Liu, K. S. X., & Park, M. (2020). A scoping review on flipped classroom approach in language education: Challenges, implications and an interaction model. *Computer Assisted Language Learning*.

Jiang, M. Y. C., Jong, M. S. Y., Lau, W. W. F., Chai, C. S., & Wu, N. (2021). Using automatic speech recognition technology to enhance EFL learners' oral language complexity in a flipped classroom. *Australasian Journal of Educational Technology, 37*(2), 110-131.

Skehan, P. (2009). Lexical performance by native and non-native speakers on language learning tasks. In B. Richards, H.M. Daller, D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application.* UK: Palgrave Macmillan.