

Student Attitudes to Using Speech-to-Text and Text-to-Speech Technology

Faculty of Cultural Development

Alexandra Burke, Margaret Yamanaka

文化創造学部

(Received December 23rd 2023)

音声テキスト変換およびテキスト音声変換テクノロジーの 使用に対する学生の態度

バーク・アレクサンドラ, 山中マーガレット

(2023年12月23日受理)

概要

この論文では、日本の教室で携帯端末を使って英語を学習することが大学生に与える動機づけの影響について、簡単な経緯を探る。学生が携帯電話に内蔵されている無料の音声読み上げ・テキスト読み上げAI技術を使い、自分の発音を聞き取ることができるように訓練した。学生たちはこの技術に対する即座の反応と、その後この技術を他の場面で活用できたかどうかについてのアンケートを試みた。アンケートに回答した15名は誰もが肯定的な反応を示した。

I. Introduction

In larger classrooms, there is limited scope to evaluate speaking performance, and except in test situations. The opportunity to give individualized, formative feedback on specific pronunciation is rare. However, this is something that students have indicated would be valuable to them and was missing from their school educational processⁱ. As a teacher of traditional music, for 40 years the author has relied heavily on the use of recording in situ. She personally applied the use of recording and listening to content as a university student, and later, as a student of Japanese from 2005. Telephones in the classroom are usually regarded as a

distraction, but can have a useful role as a tool for learning. The establishment of the Global Innovation Gateway for All (GIGA School) initiative by the Ministry of Educationⁱⁱ firmly places the role for devices in the classroom. However, there is less research on the use of the phone as a means of feedback on speaking activities.

II. A history of student usage of devices in Japanese classrooms.

In 2005, in collaboration with the English Department and the executive of Gifu Nourin High School, a new element called “Keitai Time”, was introduced to the oral communication classes.

At that point, from a student body of 900, 890 student carried cell phones. Of these, half had recording capacity of 15 seconds, but the others could receive recordings via infrared technology. The author and Japanese Teacher of English colleagues, made short, bilingual presentations of target words and expressions, which students recorded on their phones. Those who were not able to record, received the recordings via infrared transfer, a CD, or cassette. They listened to their phones for homework. The advantage was that they could study even if they were standing on a train, with no room to read a book. The impact on pronunciation and motivation was significant to the extent that the listening tests had to be increased in complexity, as students routinely scored in excess of 80 %. This information was shared with Japanese Teachers of English and Assistant Language Teachers at a Gifu Prefectural conferenceⁱⁱⁱ. The concept was adopted by other educators in Gifu and also shared as teacher-training at a Joint JACET-JALT conference^{iv}.

At a similar time, Hiroyuki Obari and colleagues, researched preferences of Japanese students for learning on devices noting that students studying on phones scored higher than those using computers^v. What is less known is the current ICT skills of students, device use and training needed^{vi}. Obari, et al, also researched the economic impact of the availability of flat-rate unlimited data packets on student preferences for learning using devices^{vii}. Access to wifi is a critical tool and many universities are now dealing with the issue of providing reliable connectivity, particularly in buildings designed for a pre-wifi era.

III. The process used at G.W.U

Students were shown to open the settings, accessibility, spoken content, and read aloud option on their phone. They selected text and listened to it. After that they were trained to find the microphone and if not done already, turn on the speech-to-text (STT) option. They were shown that it would operate on a range of applications that they routinely use. After all students understood (by working in small groups and providing peer support if needed) how to use the technology, they were encouraged to start speaking or reading into their phones.

Headphones with microphones (even low-cost headphones) make the resulting transcriptions more accurate. Reading from a text is a very good method to start with as it produces a fairly reliable result. Speaking slowly and hesitantly does not produce an accurate transcription. The best results are produced by looking away from the screen, as focusing and pausing on every word transcribed produces unusual punctuation. This experience also highlighted the importance of learning and using spoken language rules like pausing briefly for a comma and longer for a period mark. Students were able to participate in the training irrespective of whether they were going to consent to include their data and there was no impact on grade if they chose not to do the activity.

IV. Student responses analysis

Of the 34 students in the group, 17 responded and 15 consented to being included in the study. Students responded to the following prompts which were written in Japanese then

English: whether they had been aware of speech-to-text before the experience. 46 % yes / 54 % no. Had the students used the technology before the class: 33 % yes / 67 % no. After the lesson, had they used the technology again? 60 % yes / 40 % no. Of the 60 % using the technology, 5 students had incorporated it into their university study and 4 were using it for daily life activities.

Count of Have you used it again since the class (for any purpose)?

- いいえ、使ったことはありません。No. I have not used it.
- はい、大学の勉強用に。Yes. For university study.
- はい、日常生活活動に。Yes. For daily life activities.

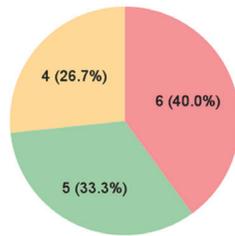


Figure 1

V. Comments from students

Students were asked “自分の言葉がデバイスの画面に表示されたのを見たときはどう思いましたか? How did you feel when you saw your spoken words appear on the device screen?” These were the responses.

発音を意識するようになった、素直にすごい技術だと思いました、とても便利だと思いました。、I was surprised by the ability to transcribe the English sentences I read aloud. and I was able to know the spelling right away.、新鮮だった。自分のイントネーションで色々な英単語が出てきて面白かった。、何も思いませんでした。I didn't think anything of it.、純粋に機械の発達を感じられてすごいと思った。使い方次第ではとても便利だと思うので上手く活用してみたいと思った。、わからない英単語があった場合に調べなくて

もすぐ変換されることに驚いた。いろいろな人に便利な良い機能だと思った。、不正確な漢字が使われることもあったけれど、比較的便利だった。、正しく認識される箇所とされない箇所があるので自分の発音が悪いのだと感じました。、すごいと思った。、すごい、きちんと表示された単語は発音ができていると感じてうれしかった。、驚きました。

VI. Potential for application in context with the GIGA School goals.

Student reactions to this functionality shows that there is potential for creating more is in the life of students, while increasing their output and engagement. A point to be considered is that MEXT recently released data that show that 8.8 % of students in public elementary and junior high schools in mainstream classes have difficulty with reading and writing without any other impairment to their academic potential. This number drops to 2.2 % in public high schools^{viii}. Voice-to-text and text-to-voice are functions which have been used for decades for people with learning difficulties to compete with neurotypical peers. The technology is embedded in Microsoft's products via Immersive Reader, because Bill Gates (and many other technology and other successful leaders) has dyslexia himself and wants barrier-free learning to become a norm. They want the next generation to grow and learn without the discrimination that they experienced^{ix}. The MEXT GIGA School and the 読書バリアフリー — initiatives^{xi} combine to create a classroom where anyone student can excel without bias and without an unfair technical or economic burden. In contrast, many other countries limit support to

those with capacity to pay in excess of ¥ 100000 for testing to prove there is a reason why they are not succeeding.

VII. Conclusion

While this is not a powerful study, it does give an indication that using STT can be empowering for students. Teachers learning to use this technology and normalizing use in the classroom may prevent students experiencing academic failure at the tertiary level^{xii}. Using the principles of education described by John Dewey of “utility, interest, experience, and integration to prepare students to participate in a changing society and life” applying accessible technology such as speech-to-text will help to achieve this^{xiii}.

- i Kobayashi D, “Sources of Speaking ESL Self-Efficacy of Japanese University Students” JALT Postconference Publication – Issue 2020.1 August 2021
- ii 文部科学省, GIGA スクール構想の実現について2019
- iii Burke A, Toyoshi M, “Keitai Time: Using your Keitai for teaching English” Gifu Prefecture JET Program Skills Development Conference 2006
- iv Scott J & Burke A: Mobile Phones as Multimedia Tools to Enhance Collaboration in the Language Classroom JACET/JALT Regional Conference, 2008
- v Kimura M, Obari H, Goda Y, Mobile Technologies and Language Learning in Japan: Learn Anywhere, Anytime Proceedings of the WorldCALL 2008 Conference, pp. 203–4
- vi 奥田由紀恵, 大学生の携帯とパソコンの利用状況 福山平成大学経営学部紀要第15号 (2019), 145–155頁
- vii Y. Goda, Y. Kogure, Y. Shimoyama, M. Kimura and H. Obari, “Survey Research on Mobile Phone Market for Mobile-Learning in Japan,” *Fifth IEEE International Conference on Wireless, Mobile, and Ubiquitous Technology in Education (wmute 2008)*, Beijing, China, 2008, pp. 194–195, doi: 10.1109 /WMUTE.2008.43 .
- viii 文部科学省, 通常の学級に在籍する特別な教育的支援を必要とする児童生徒に関する調査結果 (令和4年) について
[https://www.mext.go.jp/b_menu/houdou/2022/1421569_00005 .htm?fbclid=IwAR1h7dbJxWCBrix_wFjNoVrZ82DhgBzkQ4JMaO2DCs-1IVN2rMt3vsFjWO4](https://www.mext.go.jp/b_menu/houdou/2022/1421569_00005.htm?fbclid=IwAR1h7dbJxWCBrix_wFjNoVrZ82DhgBzkQ4JMaO2DCs-1IVN2rMt3vsFjWO4)
- ix Kim E 8 *successful tech leaders who overcame a learning disability called dyslexia* Business Insider Online Tech 2015 <https://www.businessinsider.com/tech-leaders-with-dyslexia-2015-2>
- xi 文部科学省, 誰もが読書をできる社会を目指して～読書のカタチを選べる「読書バリアフリー法」～ (啓発用リーフレット) 2019
- xii Burke A, *Reader Mode: Sweeping Away Barriers to Reading*, *The Language Teacher* 47 : 6 November 2023 pp 50–52 .
- xiii Slaughter, T. (2009). Creating a successful academic climate for urban students. *Techniques*, 16–19 . Cited in William M John Dewey in the 21st Century *Journal of Inquiry and Action in Education*, 9 (1), 2017 .