INTERNATIONAL JOURNAL OF EDUCATION, PSYCHOLOGY AND COUNSELLING (IJEPC)

www.ijepc.com

GAE

GLOBAL ACADEMIC EXCELLENCE

A CASE STUDY ON MANDARIN TEXT-TO-SPEECH

Yoke Lian Lau[1]*, Wan Hurani Osman[2], Bernadette Tobi[3], Irma Wani Othman[4]

[1] Centre for the Promotion of Knowledge and Language Learning, Universiti Malaysia Sabah, Malaysia
    Email: yokelian@ums.edu.my
[2] Centre for the Promotion of Knowledge and Language Learning, Universiti Malaysia Sabah, Malaysia
    Email: wanosman@ums.edu.my
[3] Centre for the Promotion of Knowledge and Language Learning, Universiti Malaysia Sabah, Malaysia
    Email: bernadette@ums.edu.my
[4] Centre for the Promotion of Knowledge and Language Learning, Universiti Malaysia Sabah, Malaysia
    Email: irma@ums.edu.my
*   Corresponding Author

Article Info:

Abstract:

There are three objectives of this study. First, to investigate the effectiveness of TTS in Mandarin language reading, find out the shortage of TTS, and understand the difference between machine-reading and human reading. Second, the study concludes that TTS technology is mature enough now but still needs to solve some problems in specific criteria such as conveying humanity elements, putting in more human nature, and expressing human emotions in the reading process in the future. Third, teaching mandarin reading is still necessary as the TTS system could not, instead of the importance of a teacher in reading and learning thoroughly.

Keywords:

Text-To-Speech Technology, Mandarin Language, Human Reading, Machine-Reading

Introduction

Covid-19 has forced a shakeup in the traditional learning process, with a lot of onsite teaching switching over to an online equivalent. Fortunately, IT development has prepared web services and platforms to enable successful online teaching and learning processes during the pandemic. This case study focuses on researching the functionality and effectiveness of Mandarin Text-to-Talk Speed (MTTS) using TTSMP3 (free text-to-speech and text-to-mp3).

Text-to-Speed (TTS) technology is as standard as modern speech-to-text technology. The first speech recognition toy, dating back a century to the 1920, was Radio Rex, a celluloid dog who came out from his base when he heard and recognized his name (David & Selfridge, 1962). This toy was the first hear-recognize response example in technology. During the 1950s, the Audrey system hit a goal in speech recognition technology, by enabling speech recognition technology to recognise numerical digits from 0 to 9 (Davis, 1952). Sometime later, IBM created Shoebox, a system that could recognise questions and answer them, using 16 English words, including digits from 0 to 9 along with command words such as plus, minus and total (IBM, 2021). Speech recognition has evolved since the last century, involving responses to words, numbers, actions, and commands to add convenience to daily life.

Literature Review

Both speech-to-text (STT) and text-to-speech (TTS) involve different processes. Speech-to-text, for example, involves Automatic Speech Recognition (ASR) which means a search engine hears a human voice and then converts it into readable data. More than 60 years of research have been undertaken in this field (Li, Deng, Haeb-Umback & Gong, 2015). ASR is the first process occurring when a human speaks to the speech-to-text system. A text-to-speech system with Digital Signal Processing (DSP) technology can convert speech into speech (Isewon, Oyelade & Oladipupo, 2014). Natural Language Generation (NLG) means when writing or speaking the system will predict the next word based on the structure and frequency of the language naturality (Bateman & Zock, 2012). Both speech-to-text and text-to-speech are involved in this process.

Artificial Intelligence (AI) has been around for a while and is constantly improving in capabilities. In some ways, AI has replaced the conventional role of a teacher and can enable students to do all or most of their learning online via a web service. Learning Mandarin is better done online rather than in person.

Text-to-speech (TTS) technology has been applied for various languages including Japanese (Yasuda, Wang & Yamagishi, 2021), English (Huang & Liao, 2015; Ifeanyi, Ikenna & Izunna, 2014), French (Orosanu & Jouvet, 2018), Arabic (Rebai & BenAyed, 2015), and Mandarin (Yeh, Chang & Hwang, 2013).

Text-to-speech and human speech are not exactly the same, although any gaps can be investigated from different perspectives. For example, some researchers focus on the TTS weaknesses caused by linguistic characteristics. Yasuda, Wang & Yamagishi (2021) determined under which conditions sequence-to-sequence based TTS systems would perform well when given simple input such as text or phonemes in either Japanese or English.

In addition, this study compared them to comparable deep neural network (DNN) based pipeline TTS systems using complex full context labels. The researchers enabled models generated by sequence-to-sequence based TTS methods instead of enriching them with linguistic features, in order to determine to what extent enforced models could overcome the linguistic feature limitation. Huang & Liao (2015) determined the effect on students' spelling of digital material incorporated into a TTS system. The system created its digital material using the school's Spelling Bee vocabulary list of some 300 words. 21 third-graders from a private bilingual school in Taiwan were enrolled in this study.

Rebai & BenAyed (2015) believe that TTS synthesis has been extensively studied in numerous languages, although Arabic speech synthesis is still in its infancy and is not as advanced as some other languages. Their study describes a statistical parametric approach for developing an Arabic language speech synthesis system, proposing a system for Arabic discretization to address any missing Arabic diacritic marks. Rebai & BenAyed built both discretization and synthesis systems using multilayer neural networks for in-depth learning. They also proposed many speech quality improvement methods including the use of unit-length speech databases which would be primarily implemented using concatenative-based systems. They also constructed the speech database using non-uniform unit segments to improve the speech quality in this work.

A number of researchers have been focusing on the benefits of TTS for the disabled, such as Orosanu & Jouvet (2018). Their study examined the detection of sentence modality in French applied to automatic speech-to-text transcriptions. The goal of this work was to facilitate communication with those who are deaf or hard of hearing, allowing questions to be detected and informing the user they had been detected and addressed. Hux, Wallace, Brown & Knollman-Porter (2021) have the opinion that although TTS technology could be a potential reading aid for those with aphasia, it is not yet ready to be adopted for reading. In addition, individuals with aphasia are likely to have limited TTS knowledge. Furthermore, control over user-adjustable features such as voice and selection highlighting and speech speed output settings was critical for people with aphasia. Stodden, Kelly, Takahashi, Park & Stodden (2012) will present research findings concerning the efficacy of a TTS software intervention in 2 pilot studies involving approximately 104 high school students in grades 9 to 12 with a disability which puts them at risk of being referred for special education due to their reading difficulties.

Some researchers have suggested modules for improving quality of speech. Yeh, Chang & Hweng (2013) presented a consistency analysis of a Mandarin language TTS system acoustic module as a means of enhancing speech quality, with consistency being discovered via an examination of human pronunciation processes. This can be interpreted as a high correlation of a warping curve between the spectrum and prosody within one syllable. The primary objective of the paper was to analyse the consistency of an acoustic module for speech in the Mandarin language. Several experiments have demonstrated that the warping curve between prosody and spectrum within a syllable remains consistent as long as the syllable is contained within the same word. Young, Courtad, Douglad and Chung (2018) worked on determining the efficacy of TTS on the reading comprehension and oral reading fluency (ORF) of 4 secondary school students with learning disabilities. These researchers evaluated the effectiveness of TTS

on reading outcomes using a single case A-B-A-B withdrawal design. All participants improved their reading comprehension over a period of 4 weeks of using TTS to read instructional passages. The participants' ORF results showed an increase in the number of words read per minute after each accommodation condition. Half of the participants showed increased reading scores compared to pre- and post-test achievement on the Lexile assessment.

Hallahan & Vitale (1997) discussed the software port of a mature TTS synthesis technology sold as a series of hardware products for more than a decade. The TTS quality is exceptionally high in terms of intelligibility and naturalness, and stimulates the human vocal tract using a digital formant synthesiser.

Saito, Takamichi & Saruwatari (2019) have shown that a vocoder-based system causes synthetic speech quality degradation. Despite this, Ifeanyi, Ikenna & Izunna (2014) worked for years to develop vocally interactive computers which are capable of voice and speech synthesis. An interface like this, with a computer able to generate text and deliver it in speech form, would be highly useful and beneficial. TTS synthesis is a technology enabling written text to be converted from a descriptive form to spoken language which the end-user can easily understand.

Mandarin is a tonal language (Ching, 1976; Chen, Hwang & Wang, 1996; Huang & Liao, 2015). Mandarin contains 5 tones. There is a high level with a pitch of 55, high rising with a pitch of 35, low rising with a pitch of 214, high to low falling with a pitch of 51, and neutral with a pitch of 5 (Mathews, 1931). In 1976 Votrax became the world's first Mandarin speech synthesiser (Ching, 1976; Wang, Chen, Tao & Chu,2006). This shows that the tones in Mandarin are perhaps the most challenging factor to yet be solved with technology. Mandarin has just 400 different syllables, meaning the tone of each syllable is critical in the synthesiser's ability to detect the correct word. Because of this, in the synthesis program's early stages, the pitch contour of the syllable was implemented by joining the 4 pitches (Ching, 1976; Chen, Hwang & Wang, 1996). The main issue with TTS Mandarin in the 1990s was determining which type of synthesis unit was used along with the speech synthesiser (Chen, Hwang & Wang, 1996).

Mandarin is spoken as a syllable and there are just 1300 phonetically distinguishable syllables from 411 base-syllables and 5 tones (Chen, Hwang & Wang, 1996; Wang, Chen, Tao & Chu, 2006). Although Mandarin can be streamlined to be as simple as possible, the characteristic specialisation of this language makes it an issue to solve in MTTS technology. For this reason, researchers have built up a waveform table of 411 syllables to generate prosodic information such as pitch contour, energy level, initial duration, and final syllables (Chen, Hwang & Wang, 1996).

The MTTS system requires all text to convert into phoneme sequences, known as the Pinyin system, to enable correct pronunciation in the written form of the Mandarin language (Zhang, Zhang & Zhong, 2019).

Zhang, Zhang & Zhong (2019) introduced a new method for prosodic annotation in MTTS to generate rhythmic, natural Mandarin speech patterns. The model could also predict the

prosodic boundaries between characters. This finding has been very important in the upgrade of MTTS function.

MTTS has experienced a long journey of development from the early stages which entailed focusing on finding suitable speed synthesis techniques and the selection of proper synthesis units. Next, the development moved into the stage of Pitch Synchronous Overlap and Add (PSOLA) methods and followed with the creation of more sophisticated text analysis methods with electronically-coded Chinese text input (Wang, Chen, Tao & Chu, 2006). Although the MTTS creators and researchers accomplished many improvements, there were still some problems with the MTTS, such as word segmentation becoming an essential requirement for nearly all MTTS systems and the inefficiency of identifying proper names (Wang, Chen, Tao & Chu, 2006).

Operational Definition
Text-to-speech (TTS):

According to Ifeanyi, Ikenna & Izunna (2014), TTS is a technology that converts written input to spoke output via synthetic speech generation. There are various techniques used for speech synthesis:

Straightforward voice recording and playback on demand
Speech segmented into 30-50 phonemes/fundamental linguistic units and then reconstructed into a fluent speech pattern
Approximately 400 diaphones used

Research objective and research question
The research has 3 objectives:

To investigate the effectiveness of using TTS in Mandarin language learning
To learn about the lack of TTS in Mandarin language learning
To better understand the difference between machine reading and human reading

The main research question posed is: what improvements can be achieved in TTS?

Methodology
This study includes qualitative descriptive research. It began in January 2020 and is on-going with an estimated end date of June 2022 under a Universiti Malaysia Sabah (UMS) inner grant worth RM 8000.

Procedure of Research

1. Inform research students about the research, inviting them to join as experimental subjects on a volunteer basis from January 2020.
2. Introduce varieties application in teaching and learning the Mandarin language and encourage students to use and practice it.

3. Data collection is announced at the end of the semester and students who wish to join the research register via a WhatsApp group from June 2020 to July 2020 are logged.
4. Selected students receive a note with 10 sentences chosen from the UMS Mandarin language class textbook.
5. The researchers require the students to record reading those 10 sentences in a quiet room and then return it to the researcher via email or WhatsApp.
6. Every acceptable recording will receive RM 30.

Participant Selection Criteria

All 8 of the UMS Mandarin level 2 class students volunteered for this research. This was the criteria for joining:

1. Must join the research from Mandarin level 1.
2. Must be able to read and understand the provided sentences.
3. Must not have any disabilities in performing tones and rhythms in the language.
4. Only 8 students in total are included in this study.

The researcher provided a recording to the volunteers and a total of 10 recordings were analysed.

Acceptable Reading Criteria
1. Recordings from 8 students.
2. One recording from a Mandarin language lecturer.
3. One TTSMP3 recording.

All 10 recordings had to meet these criteria:
1. The recording must be done in a quiet room.
2. The recording must be audibly clear, the pronunciation must be correct, the speed and rhythm of speech must be natural.

Selection of 10 Sentences for Reading
1. The 10 sentences are classic in the Mandarin language and may be in dialogue form.
2. The 10 sentences used were suggested by Mandarin language teachers who chose them based on their fulfilment of one or more of the following:
a. Describing Mandarin culture
b. Reflecting the Mandarin lifestyle
c. A common sentence often used in the Mandarin language

Before pre-recording, the researchers explained the meaning of each sentence to each participant and also provided the Malay language translations for each sentence. Every participant was fully briefed before recording took place.

Limitation
The study's subjects were limited in the focus area, for example Kota Kinabalu, UMS, Sabah students, and so on. Additionally, the recordings were limited to sentences appearing in the Mandarin language textbook of the UMS Mandarin language class. This resulted in just 8 recordings being used for analysis data in this study, one of which was from a native Mandarin speaker along with one from a TTSMP3 web service recording.

Tools Used
Every subject including native Mandarin language speakers used a smartphone app to create the recording. The TTS from the TTSMPS app gave a free text-to-speech recording via web service.
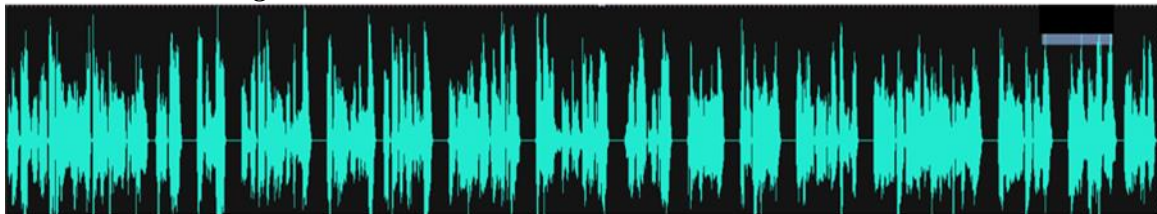
Method
Wave pad sound editor was primarily used to analyse and observe the spectrum of the recording.
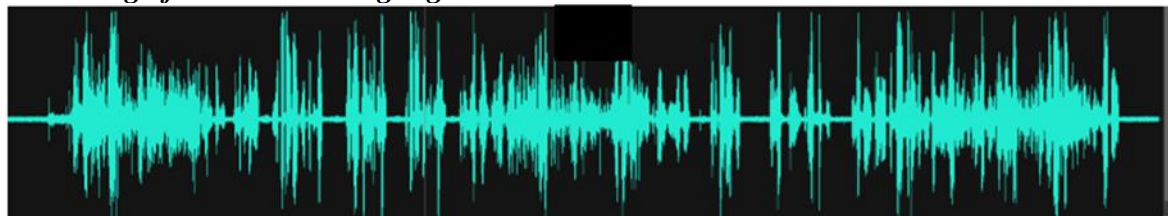
Analysis
A wave pad sound editor was used to perform a spectrum comparison of the 10 recordings. The following spectrum recordings are cut into similar length, fixed-sized boxes to make the comparison simpler.
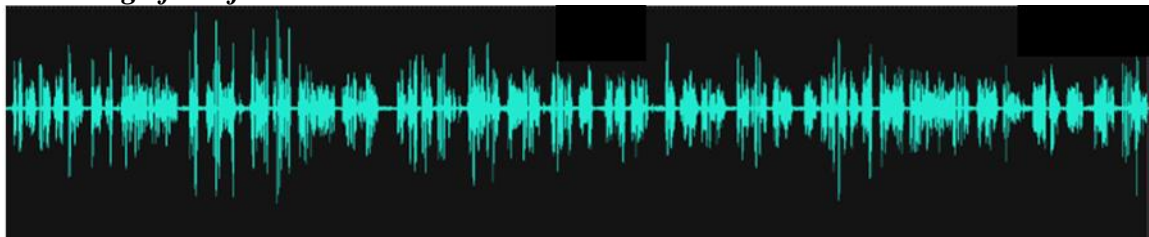
TTSMP3 Recording

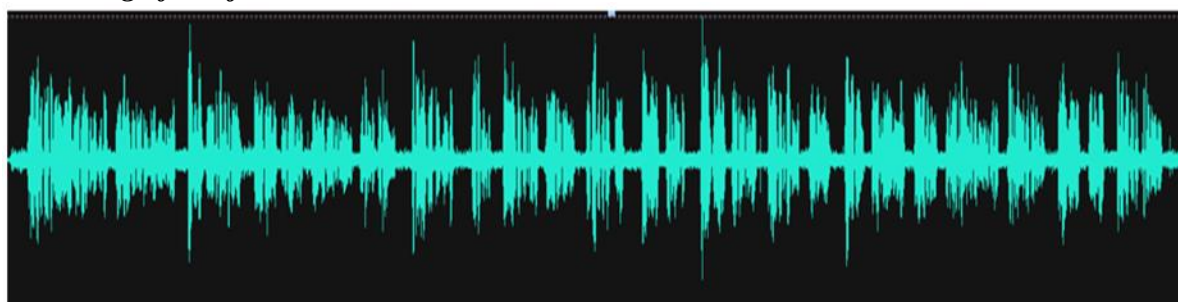*Recording of Mandarin Language Lecturer*

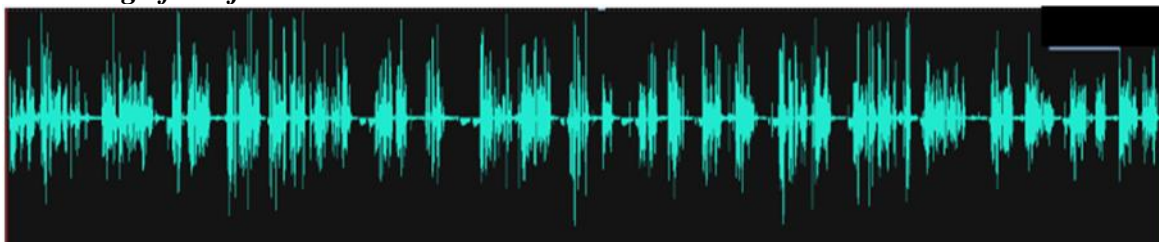*Recording of Subject 1*

Recording of Subject 2

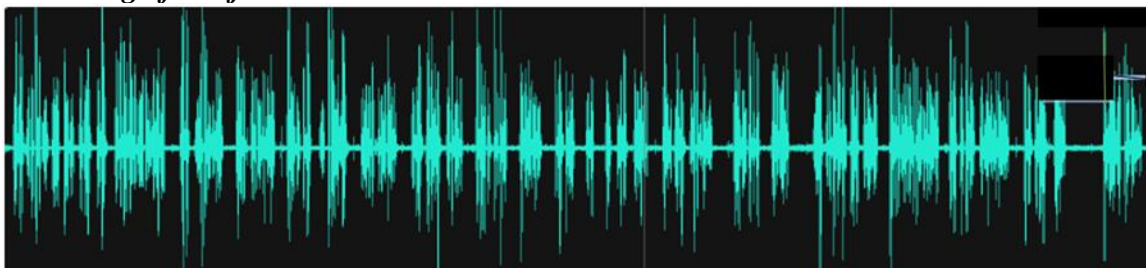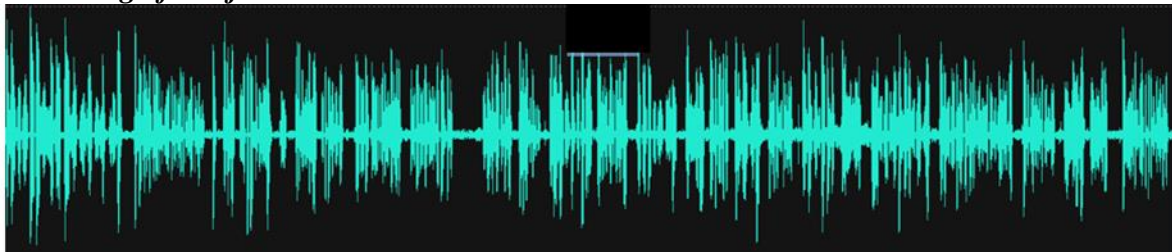*Recording of Subject 3*

Recording of Subject 4

Recording of Subject 5

Recording of Subject 6

Recording of Subject 7

Recording of Subject 8

Based on the above voicegrams/spectrograms, the horizontal axis shows the time of reading and speech while the vertical axis shows the voice frequency. The wave voice shows the amplitude (2000-2021 Keysight Technologies, 2021). With this data, the researchers were able to analyse reading tone, mood, loudness (volume), sensitivity to key/critical words, reading tone speed, and the connection between reading word to word (CBR/WTW).

Table 1: Performance of TTSMP3 System Reading and Human Reading

Subject	Reading tone	Reading mood	Loudness volume	Sensitivity to Critical words	*CBR / WTW	Reading Tone Speed
TTSMP3	In the same tone	In a calm mood	Same	No sense	Poorly connected	Average
Native speaker	Up and down	In a different mood	Different in each sentence and word	High	Connected	Fast and slow
1	Up and down	In a different mood	Different in each sentence and word	High	Connected	Fast and slow
2	Up and down	In a different mood	Different in each sentence and word	Medium	Connected	Fast and slow

3	Up and down	In a different mood	Different in each sentence and word	Low	Strong connection	Fast and slow
4	Up and down	In a different mood	Different in each sentence and word	High	Strong connection	Fast and slow
5	Up and down	In a different mood	Different in each sentence and word	High	Connected	Fast and slow
6	Up and down	In a different mood	Different in each sentence and word	High	Connected	Fast and slow
7	Up and down	In a different mood	Different in each sentence and word	Low	Poor connection	Fast and slow
8	Almost in the same tone	In a different mood	Different in each sentence and word	High	Connected	Fast and slow

*CBR/WTW: the connection between reading word to word

The TTSMP3 recording had the same tone in all sentences. Subject 8 had almost the same tone as the TTSMP3, while the remaining subjects varied the tone according to sentence meaning. The TTSMP3 mood (emotive feel) remained constant unlike the readings performed by human subjects.

The average volume (loudness) from the TTSMP3 recording remained stable while that of the human subjects varied in volume between words as well as sentences. TTSMP3 was unable to detect the keywords in each sentence and could not stress any key words throughout the 10-sentence reading process. The native speaker stressed the key words in each sentence, since this subject clearly understood the meaning, while subjects 1, 4, 5, 6 and 8 performed well with regard to key word stressing. Subjects 3 and 7 did not stress the key words so well.

TTSMP3 read well but did not seem to connect the words with regard to meaning. Subjects 3 and 4 showed the strongest connection when reading from word to word. The remaining subjects were somewhere in the middle. The TTSMP3 read everything at the same speed while the human subjects varied from fast to slow during the reading process.

Findings and Discussion

MTTS is able to produce high quality, natural-sounding synthetic speech (Wang, Chen, Tao & Chu, 2006). This is also a good indication that MTTS is developing well and is mature enough to be viable in function production.

TTSMP3 is a human-sounding form of reading which is recorded syllable by syllable and word by word. When text is converted into speech the reading machine combines all the syllables and words to form and read out the sentences. This system is not able to convey human emotions (joy, sadness, frustration, excitement, and so on) during the reading process, which is one of the weaknesses found in TTS technology.

For example, a sentence which might be obviously scary, angry, happy, funny, or sad to a human reader who understands the language will not translate into the machine-read version, since the machine is simply converting text into spoken language without adding any emotion, feeling or mood to it. As well as maintaining the same tone, the machine uses the same speed which has been programmed to complete the whole reading task instead of varying the speed based on the content of the task. Subject 8 performs in a similar way to this since this subject did not understand the meaning of the sentences, thereby reading 'like a machine'.

The volume and loudness were controlled with the TTSMP3 to remain constant, unlike the human readers. Humans tend to vary volume and loudness from word to word depending on the material being read.

The human brain is sensitive to key words while speaking. Therefore, human reading will automatically stress and emphasise what the subject perceives to be the most important words in each sentence. Machines cannot detect the theme or understand human emotion, so they are unable to do this. Some less-able Mandarin language learners also perform poorly in critical key word reading, notably subjects 3 and 7.

The human auto-reading mechanism also controls the connection when reading word-to-word. TTSMP3 and subject 7 showed a poorer performance connection, while the remaining subjects managed to connect the word-to-word element of each sentence correctly. Subjects 3 and 4 showed the strongest ability in word-to-word connection while reading.

Conclusion

The TTS Mandarin system has been able to produce natural, fluent synthetic speed sounds since the 1990s (Chen, Hwang & Wang, 1996). However, the TTS Mandarin language system and technology has some elements which could be improved to match the nature of human reading behaviour. For example, in the Mandarin language learning class, a teacher is required because the learner must learn how to read sentences with the appropriate mood, behaviour and enunciation. TTS systems are unable to convey human emotion and mood and therefore cannot 'teach' in the same way. Also, TTS systems read everything using a single tone and speed setting, although humans vary both tone and speed depending on the subject matter they are reading.

The TTS system plays an influential role in the learning to read process. It is certainly helpful for those who are weak in reading, and it is also free of charge to use.

In conclusion, the TTS system is a solid resource for self-learning for the purposes of language improvement and revision. The TTS system offers accurate pronunciation in standard sentences. Mandarin language learners can use TTS to listen to grammatically correct sentences spoke like a native Mandarin speaker. However, elements such as emotion are missing from TTS technology, setting it apart from human nature. Ultimately, the robotic nature of a TTS system can prove helpful when learning Mandarin vocabulary, grammar and sentence structure, but the additional touches of conveying emotion and mood are missing from this method of learning.

Acknowledgment

Conflict of Interest
The authors declare that there is no conflict of interest.

References
Bateman, J., & Zock, M. (2012). Chapter 15:Natural Langauge Generation. In R. Mitkov (Ed.), *The Oxford Handbook of Computational Linguistics* (1 ed., pp. 284-304). Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780199276349.013.0015

Chen., S.-H., Hwang., S.-H., & Wang, Y.-R. (1996). A mandarin text-to-speech system. *Computational Linguistics and Chinese Language Processing, 1*, 87-100.

Ching, Y. (1976). Computer synthesis of Mandarin. *ICASSP . IEEE International Conference on Acoustics, Speech, and Signal Processing*, (pp. 698-700). doi:10.1109/icassp.1976.1169963

David, E. E., & Selfridge, O. G. (1962). Eyes and Ears for Computers. *Proceedings of the IRE. 50(5)*, pp. 1093-1101. IEEE. doi:10.1109/JRPROC.1962.288011

Davis, K. H. (1952). Automatic Recognition of Spoken Digits. *J. Acoust. Soc. Am. The Journal of the Acoustical Society of America, 24*(6), 627-42.

Hallahan, W. I., & Vitale, A. J. (1997). Software text-to-speech. *International Journal of Speech Technology, 1*, 121-134.

Huang, Y.-C., & Liao, L.-C. (2015). A study of text-to-speech(TTS) in children's English learning . *Teaching English with Technology, 15*(1), 14-30.

Hux, K., Wallace, S. E., Brown, J. A., & Knollman-Porter, K. (2021). Perceptions of people with aphasia about supporting reading with text-to-speech technology: A convergent mixed methods study. *Journal of Communication Disorder, 91*, 106098.

IBM. (10 07, 2021). Retrieved 5 7, 2021, from IBM Shoebox: https://www.ibm.com/ibm/history/exhibits/specialprod1/specialprod1_7.html

Ifeanyi, N., Ikenna, O., & Izunna, O. (2014). Text-To-Speech Synthesis (TTS). *International Journal of Research in Information Technology (IJRIT), 2*(5), 154-163.

Isewon, I., Oyelade, J., & Oladipupo, O. (2014). Design and Implementation of Text To Speech Conversion for Visually Impaired People. *International Journal of Applied Information Systems (IJAIS)*, 25-30.

Li, J., Deng, L., Haeb-Umbach, R., & Gong, Y. (2015). *Robust Automatic Speech Recognition A Bridge to Practical Applications.* New York: Academic Press.

Liu, X. (2010). *New Practical Chinese Reader* (2 ed.). Beijing: Beijing Language and Culture University Press.

Mathews, R. (1931). *Chinese-English Dictionary.* China Inland Mission and Presbyterian Mission Press.

Orosanu, L., & Jouvet, D. (2018). Dedection of Sentence Modality on French Automatic Speech-to-text Transcriptions. *International Conference on Natural Language and Speech Processing, ICNLSP 2015* (pp. 38-46). Procedia Computer Science.

Rebai, I., & BenAyed, Y. (2015). Text-to-speech synthesis system with Arabic diacritic recognition system. *Computer Speech and Language, 34*, 43-60.

Saito, Y., Takamichi, S., & Saruwatari, H. (2019). Vocoder-free text-to-speech synthesis incorporating generative adversarial networks using low-/multi-frequency STFT amplitude spectra. *Computer Speech & Language, 58*, 347-363.

Stodden, R. A., K. D., Takahashi, K., Park, H. J., & Stodden, N. J. (2012). Use of text-to-speech software to improve reading skills of high school struggling readers. *International Conference on Software Development for Enhancing Accessibility and Fighting Info-exclusion (DSAI 2012). 14*, pp. 359-362. Procedia Computer Science.

Wang, R.-H., Chen, S.-H., Tao, J., & Chu, M. (2006). CHAPTER 5 MANDARIN TEXT-TO-SPEECH SYNTHESIS.

Yasuda, Y., Wang, X., & Yamagishi, J. (2021). Investigation of learning abilities on linguistic features in sequence-to-sequence text-to-speech synthesis. *Computer Speech & Language, 67*, 101183-.

Yeh, C.-Y., Chang, S.-C., & Hwang, S.-H. (2013). A consistency analysis on an acoustic module for Mandarin text-to-speech. *Speech Communication, 55*, 266-277.

Young, M. C., Courtad, C. A., Douglas, K. H., & Chung, Y.-C. (2018). The Effects of Text-to-Speech on ReadingOutcomes for Secondary Students With Learning Disabilities. *Journal of Special Education Technology*, 1-12.

Zhang, C., Zhang, S., & Zhong, H. (2019). A Prosodic Mandarin Text-to-Speech System Based. *Proceedings of APSIPA Annual Summit and Conference 2019*, (pp. 165-169).

Appendix

NAMA:

JANTINA:

UMUR:

KOD KURSUS:

AGAMA:

BAHASA ASING YANG DIAMBIL:

SAMPAI TAHAP BERAPA:

UMS KK/SANDAKAN/LABUAN:

Adakah anda seorang vegetarian: YA / TIDAK

Bayaran melalui BOOST/GRAB PAY/TNGO:

(sila bagi nombor telefon atau akaun tersebut)

Sila buat rakaman bagi 10 ayat yang tersenarai dibawah.

Pastikan anda telah terima bayaran sagu hati sebelum anda memulakan rakaman ini.

Kesemua rakaman ini akan digunakan untuk menjalankan penyelidikan psycholinguistics di mana ingin mencari tahap/level tone/frequensi/spektrum bagi bacaan ayat/perkataan yang bercanggahan dengan budaya tempatan.

Dengan menyertai penyelidikan ini sebagai responden bermakna anda telah setuju kesemua data yang dibekalkan boleh digunakan untuk menjalankan analisis dan penyelidikan di mana penyelidik berjanji akan menyimpan rakaman asal dengan sebaiknya dan tidak akan disebarkan pada awam atau pihak yang tidak berkaitan, hanya dapatan dan keputusan penyelidikan sahaja dikongsikan , semua nama dan rakaman peserta akan dirahsiakan.

Sila guna recorder untuk membuat rakaman, sila rakam dengan jelas, dan baca dengan tepat. (Sila baca juga ayat yang diterjemahkan) Kemudian sila emailkan rakaman anda kepada yokelian@ums.edu.my, terima kasih atas kerjasama dan bantuan yang diberikan. Nama File ialah Mandarin language+Nama Kamu

Bahasa Mandarin language

Kategori	ayat ambil dari buku text mandarin (Liu, 2010)	penterjemahan
1 (2) budaya	宋华一九八二年十月二十七日出生，属狗。 sòng huá yī jiǔ bā èr nián shí yuè èr shí qī rì chū shēng，shǔ gǒu。	Song Hua dilahirkan pada 27 Oktober 1982 (Tahun Anjing). (Zodiak Cina)
2 (3) budaya	A: 我很好。你和爸爸身体怎么样？ wǒ hěn hǎo。nǐ hé bà bà shēn tǐ zěn me yàng？ B: 我的身体很好，你爸爸也很好。 wǒ de shēn tǐ hěn hǎo，nǐ bà bà yě hěn hǎo。	A:Saya sangat baik, bagaimana kesihatan anda dan ayah? B: Tubuh saya sangat baik, begitu juga ayah anda. (Salam biasa yang digunakan di China)
3 (3) Tabiat /cara hidup	一斤香蕉多少钱？ yī jīn xiāng jiāo duō shǎo qián？	Berapa paun pisang itu? (unit berat yang biasa digunakan di China)
4 (1) agama	我买两瓶红葡萄酒。 wǒ mǎi liǎng píng hóng pú táo jiǔ。	Saya membeli dua botol wain merah (semasa menyambut hari jadi)
5 (1) agama	喝点儿啤酒。 hē diǎn ér pí jiǔ。	Minum bir.
6 (2) Tabiat/cara hidup	可以吸烟吗？ kě yǐ xī yān ma？	Bolehkah saya merokok?
7 (3)	我们吃寿面。 wǒ men chī shòu miàn。	Kami makan mi panjang umur (semasa menyambut hari jadi)

Cara hidup		
8 (3) Cara hidup	我想租一套房子。 wǒ xiǎng zū yī tào fáng zǐ。	Saya mahu menyewa rumah. (setelah ada teman wanita)
9 (3) Cara hidup	你愿意吃中药还是愿意吃西药？ B: 我愿意吃中药。 nǐ yuàn yì chī zhōng yào hái shì yuàn yì chī xī yào？ B: wǒ yuàn yì chī zhōng yào。	A: Adakah anda bersedia mengambil ubat tradisional Cina atau ubat Barat? B: Saya bersedia mengambil ubat tradisional Cina. (sebab ubat tradisional Cina tidak menggunakan bahan kimia tetapi secara langsung diekstrak dari bahan semula jadi)
10 (3) budaya	今年是马年，你属什么？ jīn nián shì mǎ nián，nǐ shǔ shí me？	Tahun ini adalah Tahun Kuda. Apa tahun kelahiran anda? (soalan tersirat: tanya usia seseorang)

3 kategori unsur-unsur percanggahan dengan budaya tempatan

1. Unsur-unsur yang bercanggah terus dengan ajaran agama/polisi negara/budaya tradisional
2. Tabiat/budaya/cara hidup yang ada di tempatan juga tetapi tidak popular
3. Tabiat/budaya/cara hidup yang tidak ada di tempatan tetapi tidak bercanggah dengan ajaran agama/polisi negara/budaya tradisional

Saya bersetuju menyertai penyelidikan ini, dan saya membenarkan pihak penyelidik menjalankan analisis terhadap rakaman/data yang dibekalkan bagi tujuan penyelidikan SGA 0028-2019 SAHAJA.

Tandatangan

Nama:

IC