

PENGEMBANGAN BAHAN PEMBELAJARAN BAHASA MELAYU SEBAGAI BAHASA ASING DALAM ERA AI

Hiroki Nomoto
Universiti Bahasa Asing Tokyo
nomoto@tufs.ac.jp

ABSTRAK

Makalah ini membincangkan penggunaan teknologi kecerdasan buatan (AI) dalam pengajaran bahasa Melayu sebagai bahasa asing, khususnya mengenai pembangunan bahan pembelajaran yang inovatif. Penulis memfokuskan kepada pengintegrasian teknologi penukaran teks ke pertuturan (*text-to-speech*) dalam buku teks tatabahasa penulis, iaitu *Bahasa Melayu TUFSS Bunpou* [*Bahasa Melayu TUFSS Tatabahasa*]. Perkhidmatan Text-to-Speech AI dalam Google AI telah digunakan untuk menghasilkan suara bagi hampir semua contoh yang terdapat dalam buku teks tersebut. Pendekatan ini membolehkan pelajar mendengar sebutan perkataan dan ayat dengan mudah tanpa memerlukan rakaman manual. Di samping itu, penulis juga menyediakan kelengkapan paparan makna kosa kata yang dihasilkan melalui fungsi tetikus di atas (*mouseover*), membolehkan pelajar memahami makna perkataan secara langsung tanpa merujuk kepada kamus. Kedua-dua ciri ini memudahkan pembelajaran, terutamanya untuk pelajar peringkat asas. Walaupun terdapat beberapa kekurangan seperti kesalahan sebutan dalam perkataan yang mengandungi huruf *e* dan kata pinjaman, masalah ini tidak banyak berlaku dan boleh diatasi dengan melakukan rakaman manual. Penulis menekankan potensi besar teknologi AI dalam memperkayakan bahan pembelajaran bahasa. Dengan pengintegrasian teknologi seperti ini, proses pembelajaran bahasa Melayu menjadi lebih efisien, khususnya dalam membantu pelajar mengatasi cabaran dalam menguasai kosa kata dan sebutan.

Kata kunci: bahasa Melayu sebagai bahasa asing, bahan pembelajaran, kecerdasan buatan, AI, penukaran teks ke pertuturan

PENGENALAN

Perkembangan pesat dalam bidang teknologi bahasa yang kita alami sekarang boleh membantu kita mengembangkan bahan pembelajaran bahasa asing yang bersifat inovatif. Antara teknologi yang mencapai tahap kecanggihan yang tertinggi ialah teknologi penukaran teks ke pertuturan atau *text-to-speech*, yang mana menjana suara berdasarkan input berbentuk teks. Beberapa tahun lalu, mutu hasil penjaan teknologi tersebut masih sangat rendah untuk bahasa Melayu, dengan suara janaannya merupakan jelas hasil daripada sintesis mesin. Tambahan pula, teknologi tersebut sangat terhad dan jarang ditemui di pasaran. Pada mulanya, Terjemahan Bing oleh Microsoft sahaja yang dapat membaca ayat bahasa Melayu dengan suara. Setahu penulis, Terjemahan Bing sudah memiliki fungsi ini pada tahun 2020. Kemudian, sekitar tahun 2022, Terjemahan Google juga dilengkapi dengan fungsi membaca kuat. Kini, teknologi penukaran teks ke pertuturan

ditawarkan untuk digunakan dalam pelbagai aplikasi secara meluas, melangkaui syarikat Microsoft dan Google.

Dalam makalah ini, penulis memperkenalkan cara penulis memanfaatkan teknologi tersebut dalam memperbaiki buku teks tatabahasa bahasa Melayu untuk penutur asing, iaitu *Bahasa Melayu TUFs Bunpou* [*Bahasa Melayu TUFs Tatabahasa*].¹ Khususnya, hampir kesemua contoh dalam buku teks tersebut telah ditambah suara dengan menggunakan perkhidmatan Text-to-Speech AI dalam Google Cloud.² Pada bahagian seterusnya, penulis akan memperkenalkan buku teks tersebut terlebih dahulu, diikuti dengan perbincangan mengenai bagaimana penulis telah memanfaatkan teknologi di dalamnya, termasuk teknologi penukaran teks ke pertuturan.

BUKU TEKS BAHASA MELAYU TUFs BUNPOU

Buku teks *Bahasa Melayu TUFs Bunpou* ialah sebuah buku teks dalam talian yang ditulis oleh penulis untuk digunakan dalam kelas Tatabahasa Bahasa Melayu. Penulis mula membangunkan buku teks tersebut pada tahun 2013 sebagai laman web dengan HTML. Penulis memilih bentuk laman web daripada buku kertas pada peringkat awal kerana ia membolehkan buku tersebut diakses secara percuma oleh lebih ramai pelajar, memandangkan bahan pembelajaran bahasa Melayu adalah terhad jumlahnya. Malahan harganya juga cenderung lebih mahal berbanding dengan bahan-bahan untuk bahasa lain seperti bahasa Inggeris dan bahasa Korea yang lebih tinggi permintaan pasarannya.

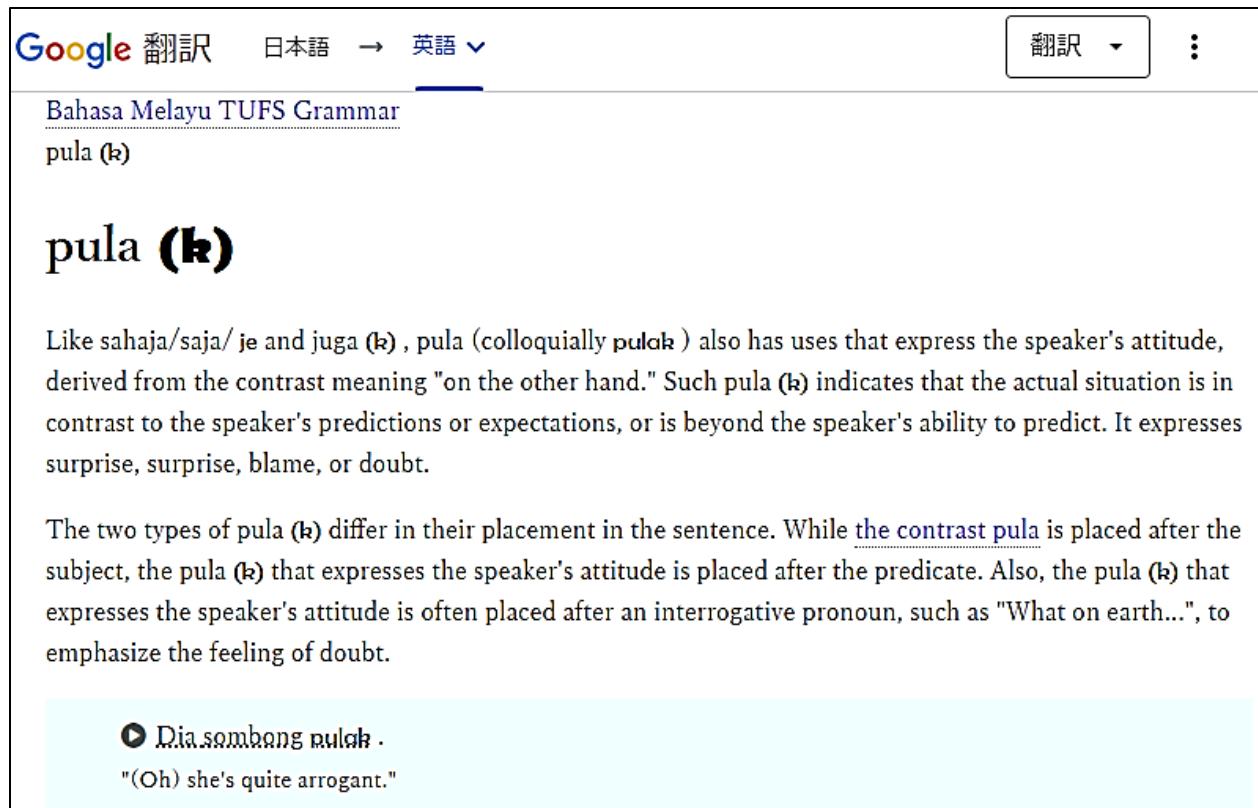
Bentuk laman web juga mempunyai pelbagai kelebihan tersendiri. Pertama, kandungan dapat diterbitkan secara beransur-ansur. Hal ini berbeza daripada bahan berbentuk buku kertas, yang hanya boleh diterbitkan setelah keseluruhan kandungannya sudah siap. Kedua, laman web mempunyai pautan yang memudahkan untuk merujuk kepada topik tatabahasa yang berkaitan. Ketiga, adalah mudah untuk memperbaiki kandungan laman web. Keempat, oleh kerana berbentuk digital, ia mudah digunakan semula. Sebagai contoh, dengan menggunakan fungsi terjemahan laman web dalam Google Translate, halaman yang ditulis dalam bahasa Jepun boleh diterjemahkan ke dalam bahasa lain (lihat GAMBAR 1). Tambahan pula, buku teks penulis diterbitkan dengan lesen CC BY 4.0.³ Hal ini bermakna pihak ketiga dibenarkan untuk membuat aplikasi dan sebagainya secara bebas dan percuma berdasarkan buku teks tersebut dengan syarat mereka memberi kredit yang sewajarnya kepada penulis.

Bentuk laman web juga mempunyai kelemahan, antaranya sukar untuk melihat keseluruhan buku. Dengan kata lain, pengguna tidak boleh memegang buku, merasai jisimnya, menyelak-nyelak lembarannya. Pengalaman fizikal sedemikian sebenarnya boleh mempengaruhi motivasi pelajar terhadap pembelajaran. Oleh itu, penulis menyusun semula buku teks dalam talian menjadi versi buku kertas tradisional dan menerbitkannya dengan judul *Mareego no Kyoukasho: Shokai Bunpou* [*Buku Teks Bahasa Melayu: Tatabahasa Komprehensif*] (Nomoto 2020; GAMBAR 2). Dalam terbitan edisi tambahan (Nomoto 2024), penulis juga mengemas kini versi asal laman web dengan menambah dua ciri baharu, iaitu paparan makna kosa kata dan suara bacaan. Dua ciri ini akan penulis bincangkan dalam bahagian selanjutnya di bawah ini.

¹ https://www.tufs.ac.jp/common/fs/ase/mal/tatabahasa_web/

² <https://cloud.google.com/text-to-speech>

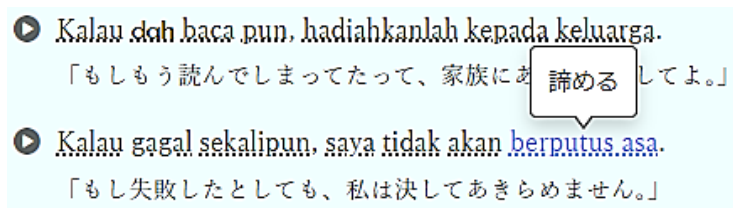
³ <https://creativecommons.org/licenses/by/4.0/deed.ms>



GAMBAR 1 Hasil terjemahan automatik ke dalam bahasa Inggeris oleh Google Translate



GAMBAR 2 Kulit buku teks tulisan penulis (Nomoto 2020)



GAMBAR 3 Contoh ciri paparan makna

DUA CIRI BAHARU DALAM EDISI TAMBAHAN

Paparan makna kosa kata

Penguasaan kosa kata merupakan antara cabaran dalam pembelajaran bahasa asing. Penulis cuba memperkenalkan kosa kata baharu secara beransur-ansur supaya jumlah kosa kata baru tidak menjadi halangan semasa pembelajaran tatabahasa. Pada masa yang sama, jumlah kosa kata baru sudah tentu akan bertambah apabila topik tatabahasa yang diajar menjadi lebih kompleks. Justeru, kebiasaannya hampir semua pelajar akan menghadapi masalah kekurangan kosa kata. Untuk meringankan beban pelajar, penulis menambah ciri paparan makna kosa kata pada laman *Bahasa Melayu TUFs Bunpou*.

Ciri paparan makna ini menggunakan fungsi tetikus di atas (*mouseover*). Apabila kursor diletakkan pada kosa kata, gelembung pertuturan yang menunjukkan makna kosa kata tersebut akan timbul di atasnya (lihat GAMBAR 3).

Penambahan makna dilakukan secara automatik menggunakan data yang juga digunakan dalam sistem bantuan bacaan Reading Tutor Bahasa Melayu (Nomoto & Kawamura 2019).⁴ Seterusnya, semua hasil penambahan automatik disemak oleh manusia. Tugas utama pada peringkat ini adalah untuk memilih makna yang sesuai dalam konteks ayat dan memadamkan makna yang tidak relevan. Contohnya, maklumat mengenai perkataan *kepada* dalam data adalah seperti berikut:

1. (人や抽象物について) ～に、～へ
2. ～にとっての

Terjemahan bagi *kepada* yang paling sesuai dalam konteks ayat pertama dalam GAMBAR 3 ialah ～に. Oleh itu, padanan lain dan keterangan penggunaan dalam kurungan telah dipadamkan.

Sara bacaan untuk contoh

Buku teks bahasa asing tidak lengkap jika tiada suara yang menunjukkan sebutan, terutamanya untuk buku-buku yang menyasarkan pelajar pada tahap asas. Dahulunya, buku teks bahasa asing dilampirkan dengan CD, tetapi kini semakin banyak buku-buku teks bahasa asing menawarkan pautan kepada fail mp3 yang boleh dimuat turun oleh pembaca. Walaupun bahan audio seperti CD dan mp3 sangat penting, proses penyediaan rakaman untuk sebuah buku teks memakan masa dan tenaga serta menelan kos yang cukup besar.

⁴ <https://malay.chuta.jp>

Teknologi penukaran teks ke pertuturan yang semakin canggih hasil daripada perkembangan AI yang pesat dapat menyelesaikan masalah tersebut. Kita tidak perlu lagi mencari penutur dan dana untuk menampung kos bayaran kepada mereka. Kita hanyalah perlu memberi kosa kata dan ayat (dokumen pun boleh) untuk dibaca kepada sistem penukaran teks ke pertuturan. Meskipun sebenarnya mutu penukaran teks ke pertuturan (masih) tidak sesempurna sebutan penutur jati, namun ia masih berguna bagi pelajar, terutamanya pada tahap asas. Oleh itu, penulis telah menambah suara bagi hampir semua contoh dalam buku teks *Bahasa Melayu TUFs Bunpou* dengan menggunakan perkhidmatan Text-to-Speech AI dalam Google Cloud.⁵ Pembaca boleh mendengar audio yang membacakan ayat contoh dengan menekan butang ►.

Semakan oleh penutur jati bahasa Melayu mendapati terdapat hasil bacaan automatik yang benar-benar tidak boleh diterima. Untuk contoh-contoh seperti itu, rakaman biasa dilakukan untuk menggantikan bacaan automatik yang dijana oleh Text-to-Speech AI. Terdapat sebanyak 22 contoh sedemikian (lihat JADUAL 1). Kebanyakan kesalahan berlaku dengan huruf *e* dan unsur asing seperti kata pinjaman. Kedua-duanya juga sukar dikenal pasti dan disebut oleh pelajar yang mempelajari bahasa Melayu sebagai bahasa asing.⁶ Pada dasarnya, huruf *e* melambangkan dua fonem, iaitu /ə/ (*e* pepet) dan /e/ (*e* taring), yang mana perlu dihafal oleh pelajar secara perkataan demi perkataan. Fonologi dan sistem ejaan kata asing serta kata pinjaman dalam bahasa Melayu merumitkan lagi situasi tersebut. Misalnya, kata pinjaman *maybe* yang sering digunakan dalam perbualan lisan dieja dengan ejaan bahasa Inggeris dan tidak dimelayukan seperti *mebi*. Huruf *e* dalam perkataan *maybe* dan *video* disebut dengan bunyi *i*. Akibatnya, huruf *e* mempunyai sebanyak tiga fonem, iaitu /ə/ dan /e/ serta /i/.

Teknologi penukaran teks ke pertuturan juga berguna apabila kita hendak menukar atau menambah contoh. Dalam situasi tersebut, jika sebelum ini, rakaman tambahan perlu dibuat. Namun kini, tugas tersebut tidak lagi diperlukan.

KESIMPULAN

Dalam makalah ini, penulis mengemukakan satu cara untuk memanfaatkan teknologi AI dalam bidang pengajaran dan pembelajaran bahasa Melayu sebagai bahasa asing, terutamanya pengembangan bahan pembelajaran. Teknologi penukaran teks ke pertuturan dibuktikan sudah cukup matang untuk digunakan sebagai pengganti rakaman audio tradisional. Penulis menggunakannya sebagai contoh dalam buku teks tatabahasa. Namun, ia juga boleh digunakan untuk bahan bacaan jika diterapkan kepada teks panjang. Bukan sahaja penyedia bahan malahan pelajar juga boleh menggunakan sendiri teknologi tersebut. Mereka dapat mengetahui sebutan perkataan, ayat atau teks tanpa rakaman yang dibuat oleh manusia lain seperti guru dan penerbit dalam bentuk CD dan sebagainya. Mereka hanya perlu menaip atau menyalin dan menampal teks yang ingin mereka ketahui sebutannya. Mutu teknologi penukaran teks ke pertuturan dan teknologi bahasa yang lain diramalkan akan meningkat lagi. Adalah diharapkan akan munculnya

⁵ Selain Google Cloud, Azure AI Speech oleh Microsoft juga menawarkan perkhidmatan penukaran teks ke pertuturan yang melibatkan bahasa Melayu.

⁶ Setahu penulis, belum ada sumber yang merakam sebutan kata pinjaman yang sebenar yang lazim digunakan orang ramai. Hal ini membawa masalah kepada pembelajar bahasa Melayu sebagai bahasa asing yang ingin menguasai bahasa sebenar, dan bukannya bahasa baku. Para pembelajar tidak mempunyai cara untuk mengetahui sebutan yang tepat. Teknologi penukaran teks ke pertuturan berkemungkinan menawarkan satu penyelesaian kepada masalah serius ini.

bahan pembelajaran bahasa Melayu sebagai bahasa asing yang lebih menarik dan inovatif yang dapat memanfaatkan sepenuhnya teknologi-teknologi terkini. Penulis berpendapat bahawa gagasan “bahasa dan teknologi” adalah kunci untuk merealisasikan harapan tersebut dengan melahirkan lebih ramai sumber manusia yang mahir dalam teknologi bahasa (Nomoto dkk. 2024).

JADUAL 1 Contoh sebutan bermasalah oleh Text-to-Speech AI

Contoh	Huruf	Masalah	Ciri
radio restoran	a	Disebut dengan <i>a</i>	Kata pinjaman
Ada apa dalam beg besar awak tu? — Ada buku teks, kamus, nota, dompet, henfon dan macam-macam lagi.	a	Disebut seperti <i>ea</i>	Kata pinjaman
Zarina, engkau dah terlupa ke?	a	Disebut dengan <i>e pepet</i>	Nama khas
kursus yang pensyarah itu ajar	a	Disebut dengan <i>e taring</i>	
mengecat pengecat pengesahan	e	Disebut dengan <i>e taring</i>	Akar bersuku kata satu
kesalahan	e	Disebut dengan <i>e taring</i>	
Tapi maybe kat sana ok kot.	e	Disebut dengan <i>e pepet</i>	Kata pinjaman
surat daripada Maserah surat yang daripada Maserah	e	Disebut dengan <i>e pepet</i>	Kata pinjaman; nama khas
Jom pergi ke Megamall minggu depan!	e	Disebut dengan <i>e pepet</i>	Kata asing; nama khas
mengomel	e	Disebut dengan <i>e pepet</i>	
berehat	e	Disebut seperti <i>i</i>	
untuk beli tiket Kita perlu isi borang ini dulu untuk beli tiket.	e	Disebut seperti <i>ye</i>	Kata pinjaman
video	e	Disebut dengan <i>e taring</i>	Kata pinjaman
kamera	a, e	Disebut sebagai <i>kamére</i>	Kata pinjaman
universiti	u	Disebut dengan <i>u</i>	Kata pinjaman
beg	b	Disebut seperti <i>m</i>	Kata pinjaman
Khamis	kh	Tiada bunyi <i>kh</i>	Kata pinjaman

RUJUKAN

Nomoto, H. (2020). Mareego no Kyoukasho: Shoukai Bunpou [Buku Teks Bahasa Melayu: Tatabahasa Komprehensif]. Tokyo: Next Publishing Authors Press.

- Nomoto, H. (2024). *Mareego no Kyoukasho: Shoukai Bunpou [Buku Teks Bahasa Melayu: Tatabahasa Komprehensif]*. (Edisi tambahan). Tokyo: Next Publishing Authors Press.
- Nomoto, H. & Kawamura, Y. 2019. Mareego Gakushuusha o Taishou Nishita Dokkai Shien Shisutemu no Kaihatsu [Pengembangan Sistem Bantuan Bacaan untuk Pembelajaran Bahasa Melayu]. *Gaikokugo Kyouiku Kenkyuu*, 22, 215–229.
- Nomoto, H., Moeljadi, D., & Farhan Athirah, A. R. (2024). Masalah Teknologi dan Isu Sosial berkaitan Penggunaan ChatGPT dalam Bahasa Melayu. *RENTAS: Jurnal Bahasa, Sastra dan Budaya*, 3(1), 1–22. <https://doi.org/10.32890/rentas2024.3.1>