White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/212543/

Version: Published Version

Article:
Handley, Zoe Louise (2024) Has artificial intelligence rendered language teaching obsolete? The Modern Language Journal. ISSN: 1540-4781

https://doi.org/10.1111/modl.12929

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Check for updates

DOI: 10.1111/modl.12929

The Modern Language Journal

COMMENTARY

Has artificial intelligence rendered language teaching obsolete?

Zoe Handley

Centre for Advanced Studies in Language and Education, Department of Education, The University of York, York, UK

Correspondence
Zoe Handley, Centre for Advanced Studies in Language and Education, Department of Education, The University of York, York YO10 5DD, UK.
Email: zoe.handley@york.ac.uk

An applied linguist with a specialization in computer-assisted language learning (CALL), I was keen to read Richard Kern's (2024, this issue) *Perspectives* column, "Twenty-first century technologies and language education: Charting a path forward." In this article, Kern identifies several possible directions language learning and teaching might take following the COVID-19 pandemic and recent technological advances. With an interest in how artificial intelligence (AI), specifically speech technologies, might be harnessed to facilitate language learning and teaching dating back to the first wave of AI (Handley, 2006, 2009; Handley & Hamel, 2005), I was particularly interested in Kern's perspectives on the impact the second wave of AI is having on language learning and teaching. As Kern points out, language and technology are intrinsically related. This has never more so been the case than for AI, the aim of which was originally to develop "machines [that can] use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves" (McCarthy et al., 1955, p. 2). As such, Kern draws our attention to the existential threat AI poses to language teaching and the need to learn a language in the first place. In the United Kingdom, like the United States, we find ourselves in a language learning crisis with the uptake of modern foreign languages (MFL) at General Certificate of Secondary Education (GCSE; i.e., primary and secondary schooling) long in decline and more than 10 university language departments closing over the last 20 years (Kenny & Barnes, 2019). I would therefore like to take this opportunity to explore in more depth whether "artificial intelligence (AI) programs like ChatGPT have all but rendered obsolete the need for personal instruction in languages other than English" (Kern, 2024, this issue, p. XX).

Perhaps it is my background in language engineering, the design of software systems that perform tasks that are dependent on processing human language (Cunningham, 1999), but I believe language educators need a more in-depth understanding of what AI is and how it works to fully appreciate the extent to which AI can replace language teachers and the need to learn a foreign language. After exploring what makes a good language teacher, I explore what AI is and the range of AI technologies that might be deployed in language teaching, before discussing the extent to which AI might replace language teachers.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *The Modern Language Journal* published by Wiley Periodicals LLC on behalf of National Federation of Modern Language Teachers Associations, Inc.

15404781, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/modl.12929 by University Library, Wiley Online Library on [15/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

WHAT MAKES A GOOD LANGUAGE TEACHER?

The professionally competent teacher

So, what makes a good language teacher? The Cognitive Action in the Classroom (COACTIV) model of teachers' professional competence (Baumert & Kunter, 2013) provides a useful starting point for this discussion. Knowledge is central to the model, which also attempts to capture the skills and personal dispositions required to be an effective teacher of any subject. Building on the seminal work of Shulman (1986), the COACTIV model claims that the effectiveness of a teacher depends not only on their knowledge of the subject (so-called content knowledge) and their general pedagogical knowledge (so-called pedagogical knowledge) but also on their "knowledge [of the subject] for teaching" (Shulman, 1986, p. 9; so-called pedagogical content knowledge). General pedagogical knowledge comprises knowledge of classroom management and organisation, as well general theories of learning and development and associated pedagogical principles. Pedagogical content knowledge is subject-specific knowledge of how easy or difficult the subject matter is to understand, including knowledge of common misconceptions and knowledge of how the subject matter can be represented to promote understanding (Shulman, 1986). In other words, pedagogical content knowledge comprises the knowledge required to plan the curriculum and individual lessons as well as select appropriate pedagogical techniques to engage learners in individual lessons. To these three dimensions, Baumert and Kunter (2013) claimed that the effectiveness of a teacher also depends on their organisational and counselling knowledge. Organisational knowledge comprises knowledge of the broader educational system as well as knowledge of the individual institution in which they are working. Counselling knowledge refers to the knowledge required to support students and their caregivers in making decisions about their education, including decisions related to careers, learning difficulties, and behavioural issues. As such, counselling knowledge is as much a skill in interpersonal communication as it is a body of knowledge. Further to this body of knowledge, the research that Baumert and Kunter (2013) drew on suggests that effective teachers have a particular personal disposition. They are confident in their ability to deploy their knowledge and skills, enthusiastic with an intrinsic interest in their subject and/or teaching, and they can cope with the demands of the role (i.e., self-regulate).

The distinctive competences of the "good language teacher"

Pedagogical content knowledge, as Shulman (1986, 1987) acknowledged, differs across subjects. Language teaching, however, is distinctive not only in terms of "the nature of the subject, the content of teaching [and] the teaching methodology" but also in terms of "teacher-learner relationships" (Borg, 2006, p. 3). Aside from the unique cultural dimension of the language curriculum, the core subject matter is the medium of instruction and the goal is to enable learners to communicate in that medium rather than to acquire knowledge of it, though equipping learners with explicit knowledge of the subject is one route to achieving that goal. As such, specific theories are required to account for second language acquisition (SLA; see, e.g., Mitchell et al., 2013, for an overview), and dedicated pedagogies have been developed (see, e.g., Larsen-Freeman & Anderson, 2011, for an overview). With the goal of enabling learners to communicate, these pedagogies require teachers to create opportunities for communication and not only facilitate interaction but also engage learners in specific patterns of interaction that support language acquisition, including negotiation of meaning (i.e., modified interaction designed to overcome communication problems; Long, 1983) and language-related episodes (i.e., dialogue about language and language use; Swain & Lapkin, 1998). This in turn requires language awareness, a form of pedagogical content knowledge specific to language teaching, that comprises an understanding of the linguistic system, typical patterns of language development, and the ability to attend to, reflect on, and talk about language (so-called metalinguistic awareness; Andrews, 2003). Moreover, the curriculum normally requires engagement in topics that are often somewhat personal in nature (Borg, 2006), and speaking in another language can be particularly anxiety provoking for some

learners (Horwitz et al., 1996). The effective language teacher therefore requires not only a distinctive form of pedagogical knowledge, but also particularly strong interpersonal or counselling skills.

WHAT IS ARTIFICIAL INTELLIGENCE?

A preliminary definition of artificial intelligence

To understand the extent to which AI might replace language teachers, it is also important to have a good understanding of what AI is. Currently, the first thing that probably comes to mind is ChatGPT. Released in November 2022, ChatGPT is a chatbot, a technology that enables a computer application to engage in conversations with end-users. It was recently made possible because of advances in natural language processing—specifically, large language models (LLMs)—and generative AI, technologies that can produce novel text and other outputs in response to user-generated prompts.

Despite claims that ChatGPT and its competitors (e.g., Bard, Bing, LlaMA, Perplexity AI, and YouChat), will revolutionise society, AI is not new. Interest in developing "thinking machines," including machines that can process human language dates to the 1950s (Turing, 1950). Generative AI is just one type of AI, AI being more broadly defined as:

> a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations or decisions influencing real or virtual environments. It uses machine and/or human-based inputs to perceive real and/or virtual environments; abstract such perceptions into models (in an automated manner e.g. with ML [machine learning] or manually); and use model inference to formulate options for information or action. AI systems are designed to operate with varying levels of autonomy. (OECD, 2019)

In other words, AI refers to a range of technologies that enable computers to perceive, learn, abstract, and reason (Launchbury, 2017). This includes technologies such as face and voice identification, voice control, autocomplete, autofill, and smart reply that already proliferate in our daily lives (Devlin et al., 2023).

AI is also far from new to language education. Grammar checkers and dictation software have long been widely available to language learners through productivity applications such as Microsoft Office (Figueredo & Varnhagen, 2006), as has dictation software (Coniam, 1999). Popular language education programs including Auralog and Dynamic English (DynEd) have also long integrated speech technology to provide learners interactive speaking practice and feedback on their pronunciation (Wachowicz & Scott, 1999), as have spoken as well as written language assessments (see, e.g., the PhonePass Test from Townshend et al., 1998, and e-Rater from Burstein et al., 1998).

A taxonomy of artificial intelligence technologies

As highlighted above, AI is not one technology. It is rather an umbrella term for a range of technologies including speech and language technologies and tools for reasoning and decision-making. The capacity to process human language is essential if AI is to replace language teachers. This capacity to understand and produce human language, which might be likened to Chomsky's "language competence," is provided in the current state-of-the-art by LLMs, statistical models of the patterns, and connections between words and phrases. These general purpose or foundation models of language in turn power applications capable of "mimicking" the human processes of reading aloud (text-to-speech synthesis), comprehension (natural language understanding), writing (text generation), speaking (speech synthesis), hearing (automatic speech recognition), listening (spoken language understanding), conversation (chatbots also referred to as spoken dialogue systems), translation (machine translation), and simultaneous interpreting (speech-to-speech translation).

15404781, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/modl.12929 by University Library, Wiley Online Library on [15/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The capacity to observe, reflect, and reason to make decisions about how to sequence content, represent it to facilitate understanding, and tailor instruction during delivery to individual learners is also essential if AI is to replace teachers. Learning analytics, defined as "the measurement, collection, analysis, and reporting of data about learners and their contexts for purposes of understanding and optimising learning and the environment in which it occurs" (Long & Siemens, 2011, para 14), refers to a class of technologies that support these functions of the teacher.

CAN ARTIFICIAL INTELLIGENCE REPLACE LANGUAGE TEACHERS?

Duolingo

So, what do these advances in AI mean for language teaching? As Kern (2024, this issue) explains, the public already appears to believe that AI-powered apps like Duolingo can replace language instruction in schools and universities, a belief it appears is shared not only by Duolingo's founder Louis von Ahn (Loewen et al., 2019), but also by some institutions that have recently announced their intention to close their language departments (Pettit, 2023). Duolingo therefore seems an appropriate starting point for this discussion. With over 80 million monthly active users worldwide,[1] Duolingo offers English language instruction to learners from more than 95 L1 backgrounds as well as courses in more than 40 different languages. While, as Kern explains, ChatGPT has recently been integrated into Duolingo Max to enable it to offer role-play activities and metalinguistic explanatory feedback, at its core, Duolingo is a

> lesson system (…) built around specific topics such as family, food, and travel; each topic introduces some grammar and cultural concepts with very limited explanations but the lessons themselves focus mostly on introducing new vocabulary and drills. The exercises offered include translation, multiple-choice word recognition questions, and spelling. Incorrect answers are handled in two ways. Some users have a 'heart' system, in which a certain number of mistakes leads to losing one out of 5 hearts (…) When all 5 hearts are lost, the user cannot practise until they recover at least some hearts. On some devices, mistakes trigger more repetition and drills and a slightly lower amount of experience or achievement gained upon completing the lesson. (Shortt et al., 2023, pp. 520–521)

Despite the apparent simplicity of this core lesson system, Duolingo is more than a first-generation rule-based intelligent tutoring system or intelligent computer-assisted language learning (ICALL; Yazdani, 1986). Some speculate that it may be supplementing examples supplied by its community of volunteers (Verweij, 2020) with automatic exercise generation (Burstein & Marcu, 2005)—another possible explanation for its infamous "silly sentences" (Leow et al., 2021). Some believe that it also employs machine translation as a reference to generate feedback (Garcia, 2013). Moreover, learning analytics are used to provide optimally spaced repetition practice (Pajak & Bicknell, 2022; Settles & Meeder, 2016) and tailor motivational messages to the individual learner (Yancey & Settles, 2020). In other words, in its core lesson system, Duolingo deploys AI to encode pedagogical content knowledge (Shulman, 1986), including to prepare lessons, sequence them, and deliver them, and maintain learners' engagement (i.e., counsel them; Baumert & Kunter, 2013).

Duolingo's effectiveness as a replacement for a human language teacher is, however, limited by its pedagogical content knowledge (Shulman, 1986), specifically its repertoire of techniques for engaging learners with content, its language awareness (Andrews, 2003), and its ability to translate that knowledge into effective feedback. While an early efficacy study commissioned by Duolingo suggested it was possible for a learner to cover the equivalent of the first semester of a college language course through just 34 hours of Duolingo use (Vesselinov & Grego, 2012), independent evaluations report mixed findings and highlight a range of limitations (Loewen et al., 2019; Shortt et al., 2023). Some learners in Loewen et al.'s (2019) study, for example, reported that they found it hard to maintain their

15404781, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/modl.12929 by University Library, Wiley Online Library on [15/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

motivation due to the repetitive nature of the activities, which focus almost exclusively on receptive language skills. In other words, Duolingo is limited by its method of language teaching and the range of techniques it uses to deliver instruction (Shulman, 1986). Some learners in Marques-Schafer and da Silva Orlando's (2018) study reported wanting more explicit grammatical feedback. In other words, there are also limitations to the way in which Duolingo represents content to learners to facilitate learning (Shulman, 1986).

Grammarly

Expanding the techniques and methods of language instruction beyond the behaviouristic drills and techniques associated with grammar translation seen in Duolingo and Memrise, other commercial language courses attempt to harness AI to provide automated feedback on writing and speaking and engage learners in simulated conversations.

With an estimated 30 million daily users (grammarly.com), Grammarly is perhaps the most well-known example of an automated writing evaluation (AWE) system that can be used to provide learners opportunities for repeated writing practice and feedback (Koltovskaia, 2020; Warschauer & Ware, 2006). AWE is an extension of automated essay scoring (AES). Widely employed in high-stakes assessment including e-Rater from Educational Testing Service (ETS), AES combines natural language processing, latent semantic analysis, and machine learning techniques to generate writing scores. Similar tools for speaking, referred to as automated speech evaluation (ASE), include SpeechRater, also from ETS, and Speechace. These systems produce scores by "training an algorithm to predict human ratings from the acoustic and linguistic features extracted from learner productions on a corpus of rated learner productions" (Handley & Wang, 2023, p. 3). In its most simple form, such an algorithm can take the form of a regression model that predicts human ratings of learner productions from "linguistic" features extracted from them using natural language processing and latent semantic analysis techniques (Zechner et al., 2009). Grammatical error detection (GED) and grammatical error correction (GEC) is carried out using a combination of natural language processing techniques including n-grams, confusion sets, and language models (Chen et al., 2024).

Evaluations of the impact of Grammarly and similar software on learners' writing, however, suggest that pedagogical support from a human tutor is necessary for students to most effectively use AWE technologies (Koltovskaia, 2020). Notably, a number of studies suggest that learners often ignore the feedback provided by these tools or are unable to use it effectively (Ranalli, 2021). In Chapelle et al.'s (2015) evaluation of criterion from ETS, for example, learners appeared to ignore half of the feedback, and only 70% of their remaining revisions resulted in improvements to their written work. As far as it is possible to establish, no similar evaluations of ASE have been conducted to date.

More recently, researchers have started exploring the potential of using ChatGPT for AWE. This research, however, suggests that while ChatGPT is able to deploy some of the pedagogical content knowledge (Shulman, 1986) you would expect of a human tutor, it is limited by its ability to counsel students (Baumert & Kunter, 2013). For example, Guo and Wang (2023) reported a comparison of the feedback provided by ChatGPT with that of teachers. They found that ChatGPT, unlike previous AWE tools, was able to provide feedback on content. Compared with teachers, ChatGPT, however, provided proportionally less feedback on content, and the feedback was less informative and less likely to engage in clarification requests. ChatGPT provided much more feedback on form than content. While this included praise, teachers felt that the feedback on form was too much and was therefore demotivating.

Rosetta Stone

Available since the late 1990s, Rosetta Stone is one of the most well-known examples of an independent-study language course that harnesses AI, specifically automatic speech recognition, to

15404781, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/modl.12929 by University Library, Wiley Online Library on [15/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

provide learners with feedback in speaking exercises and engage them in simulated conversations. Grounded in what it calls dynamic immersion, an approach similar to the natural approach (Krashen & Terrell, 1983), Rosetta Stone assumes that adults learn languages in a similar way to children. At the beginner level, having acquired receptive knowledge of vocabulary and grammar through drills in which learners are asked to select the image that matches the audio, learners move on to speaking exercises. In these exercises, the learner's task is to repeat the model word or phrase. Diagnostic feedback on pronunciation is then provided. Learners who progress beyond the beginner level also get access to the "fluency builder," which includes "conversational practice." This practice takes the form of voice-enabled branching dialogues.

The techniques for scoring and providing learners feedback on their oral proficiency in software such as Rosetta Stone are similar to those employed in the AWE tools described above (Wei et al., 2022). As for their efficacy, a study commissioned by Rosetta Stone suggested that a learner can cover the equivalent of the first semester of a college language course through 70 hours of Rosetta Stone use (Vesselinov, 2009). Independent evaluations, however, report mixed results and that learners often fail to persist with the training (Loewen et al., 2019). Despite this, CALL systems integrating automatic speech recognition to provide pronunciation practice (CAPT; e.g., Duolingo), feedback on speaking (e.g., Speechrater from ETS and Speechace), and interactive conversational practice (e.g., ELSA Speak and Speak & Improve from Cambridge) have recently proliferated. There are, however, relatively few evaluations of their impact on learners' spoken language proficiency other than those of Rosetta Stone cited above (Handley, 2018; Ngo et al., 2023). With respect to learners' experiences using these systems, it is notable that early work highlighted that the value of AI tutors equipped with automatic speech recognition might lie in the fact that they are not human and are untiring and nonjudgemental (Chiu et al., 2008). In other words, CAPT systems might complement good language teachers by enabling them to provide the intensive individualized practice required to develop speaking proficiency that they simply cannot offer themselves in their regular classes. Moreover, Computer-Assisted Pronunciation Training (CAPT) systems may be perceived to have a personal disposition (Baumert & Kunter, 2013) that is particularly well suited to counselling learners to overcome the anxiety often experienced with speaking in a foreign language (Horwitz et al., 1986).

CONCLUSION

In conclusion, AI-enabled language learning and teaching is not new. Examples of AI-powered language tutors have been available on the market since the late 1990s. Despite recent advances in AI and the speech and language technologies fundamental to enabling the development of automated language tutors, the capacities of these automated tutors remain limited compared to those of an expert language tutor. Current intelligent language tutors are not only lacking in the affective dimension and their ability to engage and counsel learners but also in the pedagogical dimension. The range of pedagogical techniques that they can employ is limited, as is their ability to represent language in a way that supports learners' understanding. Current AI language tutors at best complement expert human language tutors by offering unlimited repetitive practice and feedback on linguistic form that makes time for expert tutors to engage with knowledge and meaning, provide more creative opportunities to use language, and engage with the whole learner at a human level.

ORCID

Zoe Handley https://orcid.org/0000-0002-4732-3443

ENDNOTE

[1] https://www.statista.com/statistics/1309610/duolingo-quarterly-mau/#:~:text=Duolingo%3A%20quarterly%20global%20MAUs%202020%2D2023&text=During%20the%20third%20quarter%20of,the%20second%20quarter%20of%202022

15404781, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/modl.12929 by University Library, Wiley Online Library on [15/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

REFERENCES

Andrews, S. (2003). Teacher language awareness and the professional knowledge base of the L2 teacher. *Language Awareness*, *12*, 81–95. https://doi.org/10.1080/09658410308667068

Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, X. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Mathematics teacher education* (Vol. 8). Springer. https://doi.org/10.1007/978-1-4614-5149-5_2

Borg, S. (2006). The distinctive characteristics of foreign language teachers. *Language Teaching Research*, *10*, 3–31. https://doi.org/10.1191/1362168806lr182oa

Burstein, J., Kukich, K., Wolff, S., Lu, C., Chodorow, M., Braden-Harder, L., & Harris, M. D. (1998). Automated scoring using a hybrid feature identification technique. In C. Boitet & P. Whitelock (Eds.), Proceedings from the 36th annual meeting of the Association for Computational Linguistics (pp. 206–210). Association for Computational Linguistics. https://doi.org/10.3115/980451.980879

Burstein, J., & Marcu, D. (2005). Translation exercise assistant: Automated generation of translation exercises for native-Arabic speakers learning English. In D. Byron, A. Venkataraman, & D. Zhang (Eds.), Proceedings of HLT/EMNLP on interactive demonstrations (pp. 16–17). Association for Computational Linguistics. https://doi.org/10.3115/1225733.1225742

Chapelle, C. A., Cotos, E., & Lee, J. (2015). Validity arguments for diagnostic assessment using automated writing evaluation. *Language Testing*, *32*, 385–405. https://doi.org/10.1177/0265532214565386

Chen, B., Bao, L., Zhang, R., Zhang, J., Liu, F., Wang, S., & Li, M. (2024). A multi-strategy computer-assisted EFL writing learning system with learning incorporated and its effects on learning: A writing feedback perspective. *Journal of Educational Computing Research*, *61*, 60–102. https://doi.org/10.1177/07356331231189294

Chiu, T. L., Liou, H. C., & Yeh, Y. (2008). A study of web-based oral activities enhanced by automatic speech recognition for EFL college learning. *Computer Assisted Language Learning*, *20*, 209–233. https://doi.org/10.1080/09588220701489374

Coniam, D. (1999). Voice recognition software accuracy with second language speakers of English. *System*, *27*, 49–64. https://doi.org/10.1016/s0346-251x(98)00049-9

Cunningham, H. (1999). A definition and short history of language engineering. *Natural Language Engineering*, *5*, 1–16. https://doi.org/10.1017/s1351324999002144

Devlin, H., Cousins, R., & Amitrano, A. (2023, 25 October). A day in the life of AI. The Guardian. https://www.theguardian.com/technology/ng-interactive/2023/oct/25/a-day-in-the-life-of-ai

Figueredo, L., & Varnhagen, C. K. (2006). Spelling and grammar checkers: Are they intrusive? *British Journal of Educational Technology*, *37*, 721–732. https://doi.org/10.1111/j.1467-8535.2006.00562.x

Garcia, I. (2013). Learning a language for free while translating the web. Does Duolingo work? *International Journal of English Linguistics*, *3*, 19–25. https://doi.org/10.5539/ijel.v3n1p19

Guo, K., & Wang, D. (2023). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies (2023)*, 1–29. https://doi.org/10.1007/s10639-023-12146-0

Handley, Z. (2006). Is text-to-speech synthesis ready for use in computer-assisted language learning (CALL)? [Unpublished doctoral dissertation]. The University of Manchester.

Handley, Z. (2009). Is text-to-speech synthesis ready for use in computer-assisted language learning? *Speech Communication*, *51*, 906–919. https://doi.org/10.1016/j.specom.2008.12.004

Handley, Z. (2018). Replication research in computer-assisted language learning: Replication of Neri et al. (2008) and Satar & Özdener (2008). *Language Teaching*, *51*, 417–429. https://doi.org/10.1017/s0261444817000040

Handley, Z., & Hamel, M. J. (2005). Establishing a methodology for benchmarking speech synthesis for computer-assisted language learning (CALL). *Language Learning & Technology*, *9*, 99–120. https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/cb8265da-19c9-4708-966d-af27efb74137/content

Handley, Z., & Wang, H. (2023). What do the measures of utterance fluency employed in Automatic Speech Evaluation (ASE) tell us about oral proficiency? *Language Assessment Quarterly*, *21*, 3–32. https://doi.org/10.1080/15434303.2023.2283839

Horwitz, E. K. (1996). Even teachers get the blues: Recognizing and alleviating language teachers' feelings of foreign language anxiety. *Foreign Language Annals*, *29*(3), 365–372.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, *70*, 125–132. https://doi.org/10.2307/327317

Kenny, N., & Barnes, H. (2019). Language learning: Turning a crisis into an opportunity. *British Academy Review*, *35*, 12–13.

Kern, R. (2024). Twenty-first century technologies and language education: Charting a path forward. *Modern Language Journal*, *108*, XX–XX.

Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, *44*, 100450. https://doi.org/10.1016/j.asw.2020.100450

Krashen, S. D., & Terrell, T. (1983). *Natural approach*. Pergamon.

Larsen-Freeman, D., & Anderson, M. (2011). *Techniques and principles in language teaching* (3rd ed.). Oxford University Press.

Launchbury, J. (2017). A DARPA perspective on artificial intelligence. DARPA. https://www.darpa.mil/attachments/AIFull.pdf

Leow, J., Bialostozky, J., & Blanco, C. (2021, 3 March). How silly sentences can help you learn. Duolingo Blog. https://blog.duolingo.com/how-silly-sentences-can-help-you-learn/

15404781, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/modl.12929 by University Library, Wiley Online Library on [15/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Loewen, S., Crowther, D., Isbell, D. R., Kim, K. M., Maloney, J., Miller, Z. F., & Rawal, H. (2019). Mobile-assisted language learning: A Duolingo case study. *ReCALL*, *31*, 293–311. https://doi.org/10.1017/s0958344019000065

Long, M. H. (1983). Native speaker/non-native speaker conversation and negotiation of comprehensible input. *Applied Linguistics*, *4*, 126–141. https://doi.org/10.1093/applin/4.2.126

Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, *46*, 30.

Marques-Schafer, G., & daSilva Orlando, A. A. (2018). Languages learning conceptions and Duolingo: A critical analysis on its proposals and learners' experiences. *Texto livre-linguagem e Tecnologia*, *11*, 228–251.

McCarthy, J., Minsky, M. L, Rochester, N., & Shannon, C. E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence. http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html

Mitchell, R., Myles, F., & Marsden, E. (2013). *Second language learning theories*. Routledge.

Ngo, T. T. N., Chen, H. H. J., & Lai, K. K. W. (2023). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, *36*, 4–21. https://doi.org/10.1017/s0958344023000113

OECD. (2019). Artificial intelligence in society. OECD Library. https://www.oecd-ilibrary.org/science-and-technology/artificial-intelligence-in-society_eedfee77-en

Pajak, B., & Bicknell, K. (2022, 14 September). At Duolingo, humans and AI work together to create a high-quality learning experience. Duolingo Blog. https://blog.duolingo.com/how-duolingo-experts-work-with-ai/

Pettit, E. (2023, 18 August). Scholars see dangerous precedent in West Virginia U.'s plan to cut foreign languages. The Chronicle of Higher Education. https://www.chronicle.com/article/scholars-see-dangerous-precedent-in-west-virginia-u-s-plan-to-cut-foreign-languages

Ranalli, J. (2021). L2 student engagement with automated feedback on writing: Potential for learning and issues of trust. *Journal of Second Language Writing*, *52*, 100816. https://doi.org/10.1016/j.jslw.2021.100816

Settles, B., & Meeder, B. (2016). A trainable spaced repetition model for language learning. In K. Erk & N. A. Smith (Eds.), Proceedings of the 54th annual meeting of the association for computational linguistics (Vol. 1, pp. 1848–1858). Berlin, Germany. https://doi.org/10.18653/v1/p16-1174

Shortt, M., Tilak, S., Kuznetcova, I., Martens, B., & Akinkuolie, B. (2023). Gamification in mobile-assisted language learning: A systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning*, *36*, 517–554. https://doi.org/10.1080/09588221.2021.1933540

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*, 4–14. https://doi.org/10.2307/1175860

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, *57*, 1–23. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal*, *82*, 320–337. https://doi.org/10.1111/j.1540-4781.1998.tb01209.x

Townshend, B., Bernstein, J., Todic, O., & Warren, E. (1998). Estimation of spoken language proficiency. In R. Carlson, K. Dunger, B. Granström, & A. M. Öster (Eds.), Proceedings of the ESCA Workshop STiLL "Speech Technology in Language Learning" (pp. 179–182). ESCA.

Turing, A. M. (1950). Computer machinery and intelligence. *Mind*, *59*, 433–460.

Verweij, R. (2020). Automated exercise generation in mobile language learning. Bard College Senior Projects, 297. https://digitalcommons.bard.edu/senproj_s2020/297

Vesselinov, R. (2009). Measuring the effectiveness of Rosetta Stone: Final report. http://vesselinov.com/Effectiveness%20Study%202008_RS.pdf

Vesselinov, R., & Grego, J. (2012). Duolingo effectiveness study: Final report. https://www.languagezen.com/pt/about/english/Duolingo_Efficacy_Study.pdf

Wachowicz, K. A., & Scott, B. (1999). Software that listens: It's not a question of whether, it's a question of how. *CALICO Journal*, *16*, 253–276. https://doi.org/10.1558/cj.v16i3.253-276

Warschauer, M., & Ware, P. (2006). Automated writing evaluation: Defining the classroom research agenda. *Language Teaching Research*, *10*, 157–180. https://doi.org/10.1191/1362168806lr190oa

Wei, X., Cucchiarini, C., van Hout, R. W. N. M., & Strik, H. (2022). Automatic speech recognition and pronunciation error detection of Dutch non-native speech: Cumulating speech resources in a pluricentric language. *Speech Communication*, *144*, 1–9. https://doi.org/10.1016/j.specom.2022.08.004

Yancey, K., & Settles, B. (2020, 3 September). "Hi, it's Duo": Meet the AI behind the meme. Duolingo Blog. https://blog.duolingo.com/hi-its-duo-the-ai-behind-the-meme/

Yazdani, M. (1986). Intelligent tutoring systems: An overview. *Expert Systems*, *3*, 154–163. https://doi.org/10.1111/j.1468-0394.1986.tb00488.x

Zechner, K., Higgins, D., Xi, X., & Williamson, D. M. (2009). Automatic scoring of non-native spontaneous speech in tests of spoken English. *Speech Communication*, *51*, 883–895. https://doi.org/10.1016/j.specom.2009.04.009

How to cite this article: Handley, Z. (2024). Has artificial intelligence rendered language teaching obsolete? *Modern Language Journal*, 1–8. https://doi.org/10.1111/modl.12929

15404781, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/modl.12929 by University Library, Wiley Online Library on [15/05/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License