

# Voice

## AUDIO · VIDEO

Updated on December 13, 2024

### Natural Versus Synthesized Voice

Instructors have the option of lecturing with the human voice or a computer-generated voice. In a study involving a lecture on lightning formation, students learned better from the recorded human voice than the Microsoft text-to-speech software (Atkinson et al., 2005). This finding aligns with the **voice principle**, which asserts that the human voice is superior to a computer-generated voice for lecturing (Fiorella & Mayer, 2021). However, improvements to voice synthesis technology have made computer-generated voices similar to or better than human voices for learning (Craig & Schroeder, 2017).

A Google search for text-to-speech applications can bring up many choices. Among these choices, Textmagic ([freetools.textmagic.com/text-to-speech](https://freetools.textmagic.com/text-to-speech)) and NaturalReader ([naturalreaders.com/online](https://naturalreaders.com/online)) came up early in the search and are free to use without logging in. To use them, select the desired voice from the dropdown menu, enter text into the text field, and press the “play” button to have the input text read aloud (Figure 1).

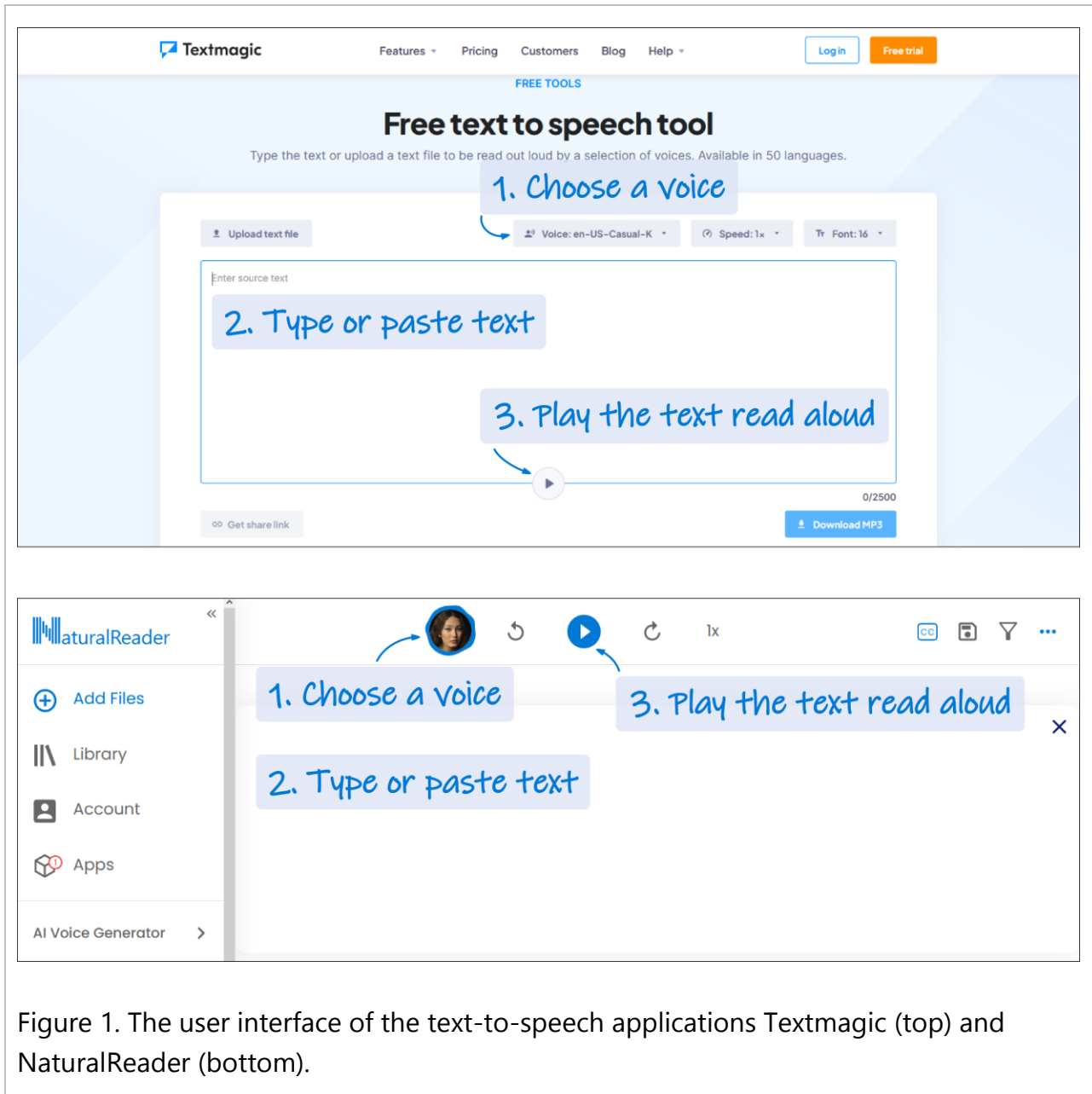














Figure 1. The user interface of the text-to-speech applications Textmagic (top) and NaturalReader (bottom).

Even though synthetic voices have evolved from sounding robotic to more life-like, lecture content that has large amounts of heteronyms or technical words may exceed the capabilities of voice synthesizers and, thus, require the course developer’s subject matter expertise to speak these words (Table 1).

Table 1. A comparison of output from free, online voice synthesizers that do not require an account to use. The Textmagic voice is "en-US-Casual-K", and the NaturalReader voice is "Jane", in English (US). The synthesized voices are compared to readings of the same texts by course developer Jonathan.

Written input	Jonathan	Textmagic "en-US-Casual-K"	NaturalReader "Jane"
Lead(II) nitrate in water can lead to environmental hazards.		 <i>Remark: Awkward pauses when saying "lead (II) nitrate".</i>	 <i>Remark: The ionic charge "(II)" is omitted.</i>
"Lead" is a heteronym. 1. Pronounced as "lead". To initiate or to be in a position of initiative or advantage. 2. Pronounced as "led". (Science) A metal.			
The unionized group of teachers believe that the side chains of acidic amino acids are unionized below the isoelectric point.		 <i>Remark: Mispronounced heteronym.</i>	 <i>Remark: Mispronounced heteronym.</i>
"Unionized" is a heteronym. 1. Pronounced as "union-ized". (Social science) Workers belonging to an organization that promotes their interests. 2. Pronounced as "un-ionized". (Chemistry) Does not carry an electrical charge.			
carbocation		 <i>Remark: Mispronounced technical word.</i>	
"Carbocation", pronounced as "carbo-cat-ion", is a technical word in organic chemistry referring to an organic molecule with a positively charged carbon atom.			
cyclooctatetraene		 <i>Remark: Mispronounced</i>	 <i>Remark: Mispronounced</i>

technical word. technical word.

“Cyclooctatetraene” is a cyclic molecule with alternating single and double bonds and has the formula C<sub>8</sub>H<sub>8</sub>. The word is pronounced as “cyclo-octa-tetra-eeen”. The “oo” and “ae” are not diphthongs.

Tk’emlúps te Secwépemc



*Remark:* He tries his best according to the pronunciation by Thompson Rivers University (n.d.).

*Remark:* Computational difficulty.

*Remark:* Computational difficulty.

“Tk’emlúps te Secwépemc” means “the people of the confluence” and is an Indigenous group of people in British Columbia who speak Interior-Salish Secwepemc (Tkemlúps te Secwépemc, n.d.).

Moreover, the synthesized voice may convey an emotion that does not match the spoken message, which may confuse the listener (Table 2). Also, the mismatch in emotion may make it difficult to use vocal cues to emphasize key words.

Table 2. Voice synthesizers may have limited ability in expressing appropriate emotion.

Written input	Jonathan	Textmagic “en-US-Casual-K”	NaturalReader “Jane”
Stop them! They’re stealing my car!			

## Foreign Accents

Whether lecturing with the human voice or a synthesized voice, a point to consider is the accent. One study found that lecturing in an unfamiliar accent affected learning negatively (Chan et al., 2020). In the study, college students who were native speakers of American English listened to a slideshow about lightning formation. Students learned poorly when the slideshow was narrated by someone whose native language was Cantonese and spoke with a foreign accent compared to that of an American who speaks Standard American English. The potential remedies are to display closed

captioning when the lecturer speaks with a foreign accent (Chan et al., 2020) or to lecture in a synthesized voice that matches the students' accent.

Although the study (Chan et al., 2020) provides valuable insight, it can have problematic implications. The issue of discriminatory hiring practice arises when we consider whether we should only hire instructors who speak English with a Canadian accent for the presumption of better learning.

Another interpretation of the foreign accent study is that the lecture is best delivered in a manner of speaking familiar to the students. Hence, if most of the students in a class speak a native language other than English, then perhaps we should hire instructors with a matching native, or faked, accent to maximize the numbers of students who can learn effectively. Such a resolution would seem rather awkward.

When hiring, there could be human rights issues with discriminating based on how someone speaks. Discrimination according to language proficiency may be acceptable (BC Human Rights Tribunal, n.d.). However, discriminating based on someone's accent may be interpreted as racial discrimination (BC Human Rights Tribunal, n.d.; Ontario Human Rights Commission, 2009).

## Summary

- Even though artificial intelligence has made tremendous progress in voice synthesis, the technology is imperfect.
- The human voice, rather than a synthesized voice, is preferred for lecturing.
- We all speak with some sort of accent, and speaking accent may be a contentious issue when hiring lecturers.

## Media Attributions

All figures are screenshots taken and used under Fair Dealing guidelines.

## References

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice.

- Contemporary Educational Psychology*, 30(1), 117–139.  
<https://doi.org/10.1016/j.cedpsych.2004.07.001>
- BC Human Rights Tribunal. (n.d.). *Leading cases: Protected characteristics*. Retrieved October 19, 2024, from <https://www.bchrt.bc.ca/law-library/leading-cases/protected-characteristics/#race>
- Chan, K. Y., Lyons, C., Kon, L. L., Stine, K., Manley, M., & Crossley, A. (2020). Effect of on-screen text on multimedia learning with native and foreign-accented narration. *Learning and Instruction*, 67, Article 101305.  
<https://doi.org/10.1016/j.learninstruc.2020.101305>
- Craig, S. D., & Schroeder, N. L. (2017). Reconsidering the voice effect when learning from a virtual human. *Computers & Education*, 114, 193–205.  
<https://doi.org/10.1016/j.compedu.2017.07.003>
- Fiorella, L., & Mayer, R. E. (2021). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (pp. 277–285). Cambridge University Press.
- Ontario Human Rights Commission. (2009). *Policy on discrimination and language*.  
[https://www.ohrc.on.ca/sites/default/files/attachments/Policy\\_on\\_discrimination\\_and\\_language.pdf](https://www.ohrc.on.ca/sites/default/files/attachments/Policy_on_discrimination_and_language.pdf)
- Thompson Rivers University. (n.d.). Retrieved October 19, 2024, from [https://www.tru.ca/\\_\\_shared/assets/tkemplups-te-secwepemc46997.mp3](https://www.tru.ca/__shared/assets/tkemplups-te-secwepemc46997.mp3)
- Tk'emlúps te Secwépemc. (n.d.). *Our land*. Retrieved October 19, 2024, from <https://tkemplups.ca/profile/history/our-land>

#Audio  
#Video

#Chemistry  
#Voice principle

#History

#Organic chemistry