Microsoft

# Efficacy of text-to-speech technology in reducing barriers to inclusive reading in higher education

Cathal O'Connor
Principal Data Scientist
Microsoft and UCD (part-time PhD)


Supervisor: Dr. Anthony Cronin

UCD DUBLIN

SFI Centre for Research Training in Foundations of Data Science

# Topics

1. State of Higher Education on Accessibility, Equity, and Inclusion
2. Challenges of Maths for students with dyslexia
3. Inclusively Designed Technology
4. Research related to Microsoft Immersive Reader and Maths Assistant
5. Future Research – Neural Speech Synthesis

# Empower every student, educator and institution to achieve more.

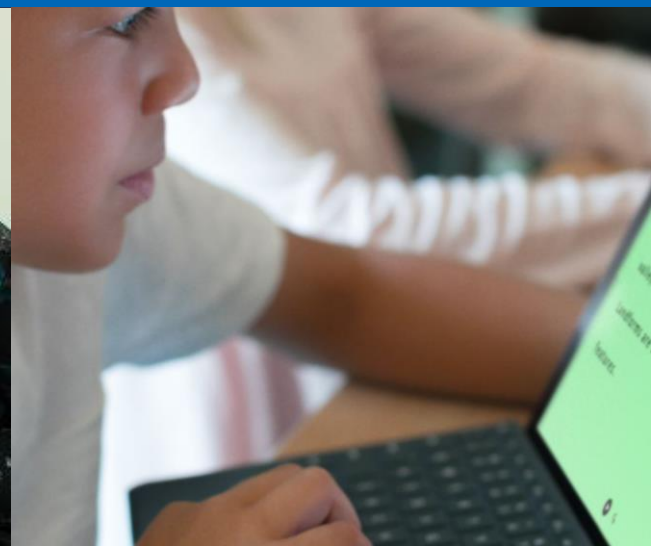
# Increased Demand for Accessibility

560% Increase in Immersive Reader | Call volumes up 200% | 30X increase in Teams captioning


Accessible Events, Climbing Toddlers and Barking Dogs

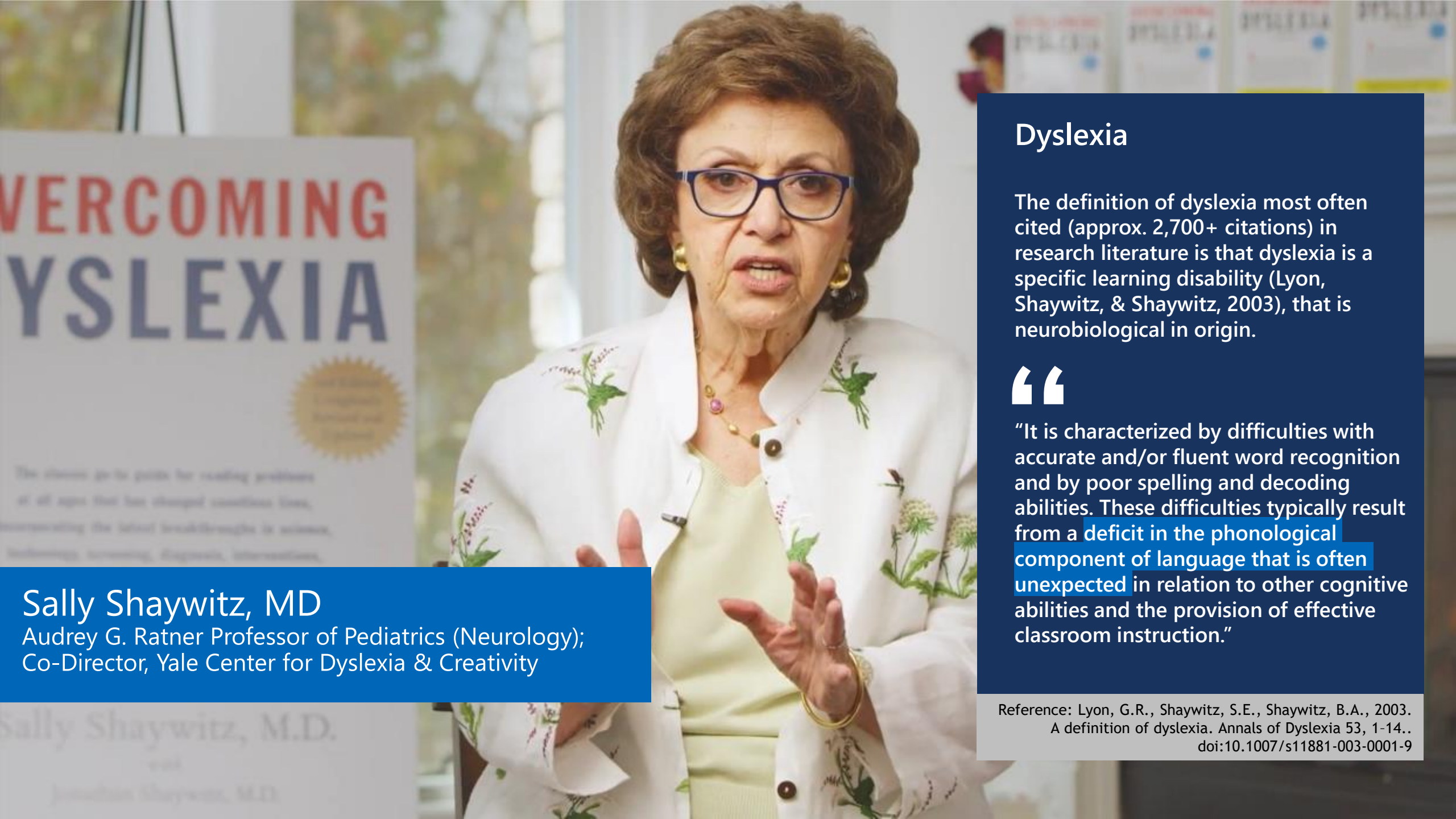
What the disability community can teach us about working remotely

Getting into Gaming: Top Tips for Accessible Gameplay



Tips for your at-home students with disabilities

[Access more information at our Accessibility Blog]




## Sally Shaywitz, MD

Audrey G. Ratner Professor of Pediatrics (Neurology);
Co-Director, Yale Center for Dyslexia & Creativity

## Dyslexia

The definition of dyslexia most often cited (approx. 2,700+ citations) in research literature is that dyslexia is a specific learning disability (Lyon, Shaywitz, & Shaywitz, 2003), that is neurobiological in origin.

> "It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction."

Reference: Lyon, G.R., Shaywitz, S.E., Shaywitz, B.A., 2003. A definition of dyslexia. Annals of Dyslexia 53, 1-14.. doi:10.1007/s11881-003-0001-9

# Dyslexia is common...

dyslexia association of ireland

**1 in 10**
people in Ireland have dyslexia, many of whom need assistive technology

But up to
**80%**
are not spotted in school

and
**dyslexia can have mental health implications**

Up to
**1 in 5**
people in the US have dyslexia

**Shame**

**Embarrassment**

**Low Expectations**

# Challenges of maths for students with dyslexia

## OUTCOMES AND EFFECTS

1. **Language of maths**
2. **Sequencing**
3. **Orientation**
4. **Working memory**

> A lack of **confidence** in their own maths ability can increase the **anxiety** and challenges of maths for students with dyslexia.

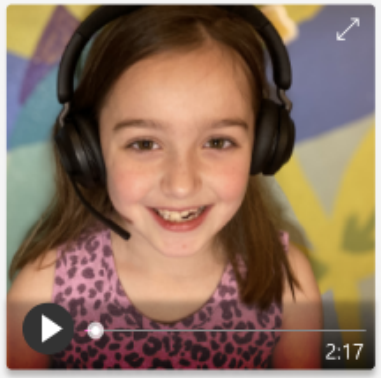
* Source: Dyslexia Association of Ireland.

Search

Activity
Chat
Teams
Assignments
Calendar
Files
...
Store
Help

Reading Progress

Close




Auto-Detect PREVIEW

Pronunciation sensitivity
Low

# Christie Cline

Attempts: 2 Level: B Words: 207 Insights

**41**
Correct Words per Minute

**83%**
Accuracy Rate

5 Mispronunciations
2 Omissions
1 Self-corrections
1 Insertions
0 Repetitions

**5**
Practiced Words

The study of Earth's landforms is called physical geography. Landforms can be mountains and valleys. They can also be glaciers, lakes, or rivers. Landforms are sometimes called physical features. It is important for students to know about the physical geography of Earth.

The seasons, the atmosphere and all the natural processes of Earth affect where people are able to live. Geography is one of a combination of factors that people use to decide where they want to live.

The physical features of a region are often rich in resources. Within a nation, mountain ranges become natural borders for settlement areas. In the U.S., major mountain ranges are the Sierra Nevada, the Rocky Mountains, and the Appalachians.

Fresh water sources also influence where people settle. People need water

Scarlet Laurie

Student work
Not turned in View history

Geography of the Earth ...

Feedback

Great job finishing all your practice words this week Christie! Just remember|

Points
/ 100

Return

# Inclusively designed
## for the benefit of all

Inclusivity and accessibility checks built-in across products and services

Each student and educator can customise the type and level of support required

Multiple ways for students to create, engage, and participate regardless of ability

**Microsoft Teams Live Captions:** View live captions and subtitles in up to six languages

**Windows Ease of Access Center:** Enable students and teachers to configure devices to meet their specific needs

**Microsoft Editor:** Bring out a student's best writer in more than 20 languages with the help of AI

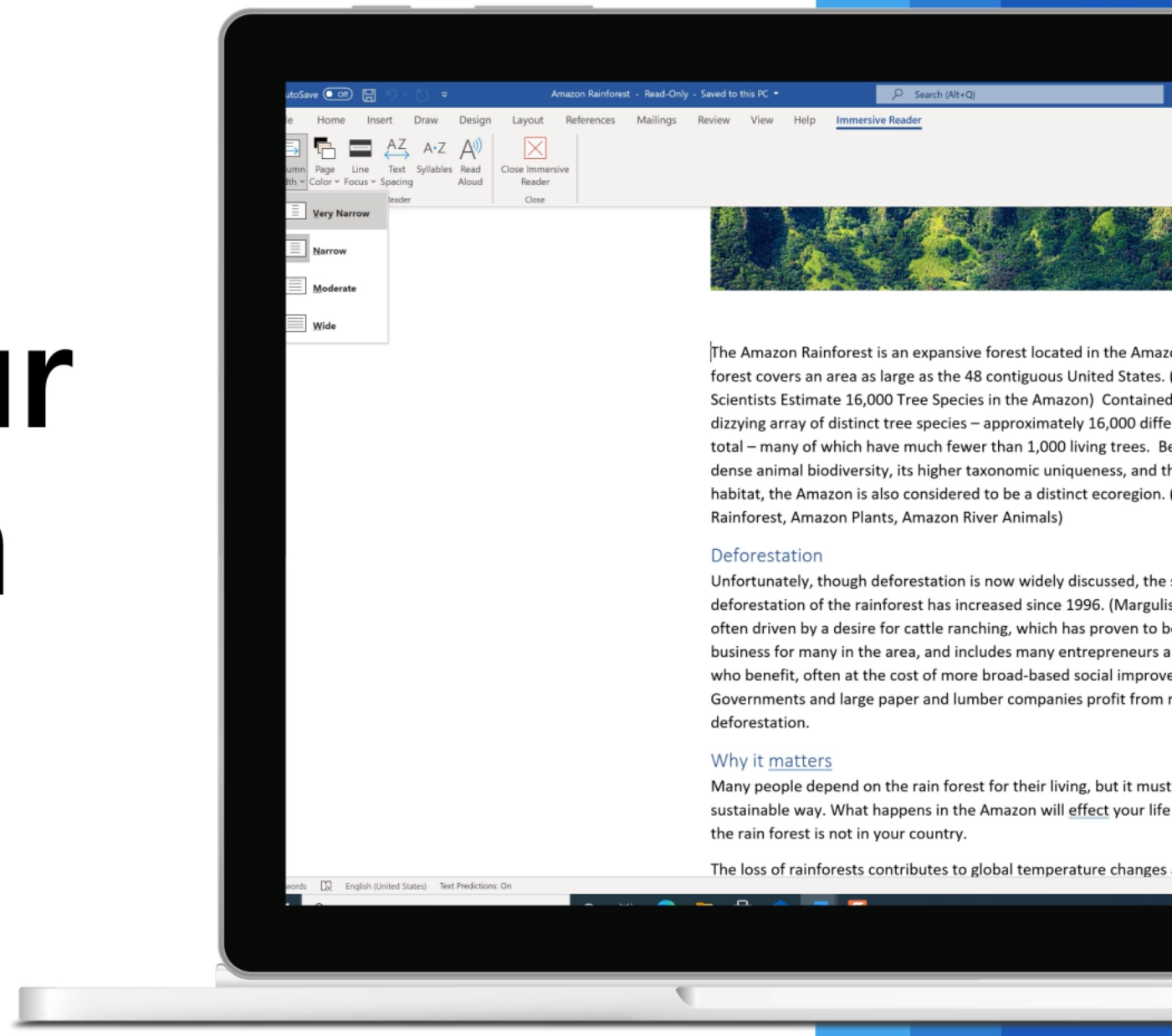
***Immersive Reader:** Improve reading for students regardless of age or ability to build comprehension and encourage independent learning

**Dictation:** Use dictation to convert spoken words into text anywhere on your PC with built-in speech recognition

****Maths Assistant:** Solve any maths equation, or display the step-by-step instructions guiding students to reach the solution on their own

Microsoft Word

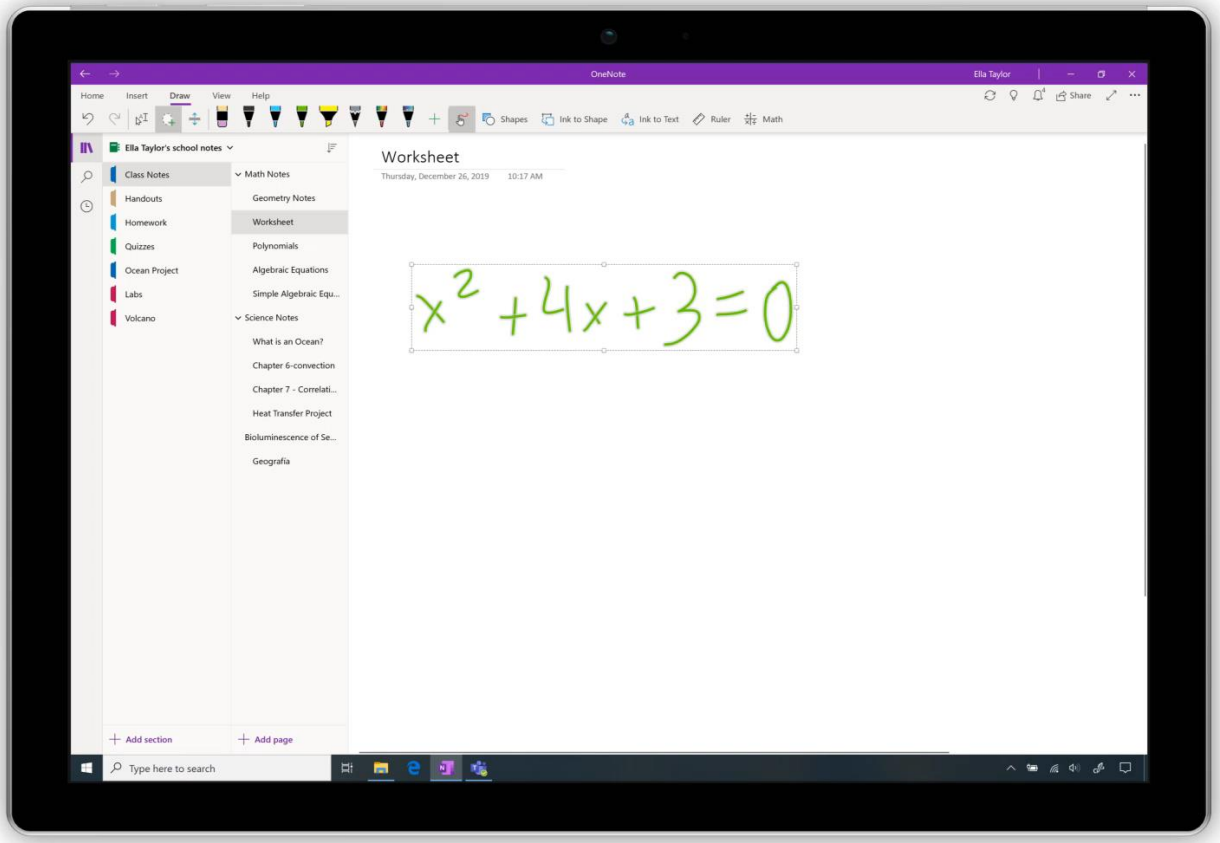
# Improve your reading with Immersive Reader



# Maths Assistant in OneNote

- Personalized experience
- Master important maths skills
- Equation read aloud

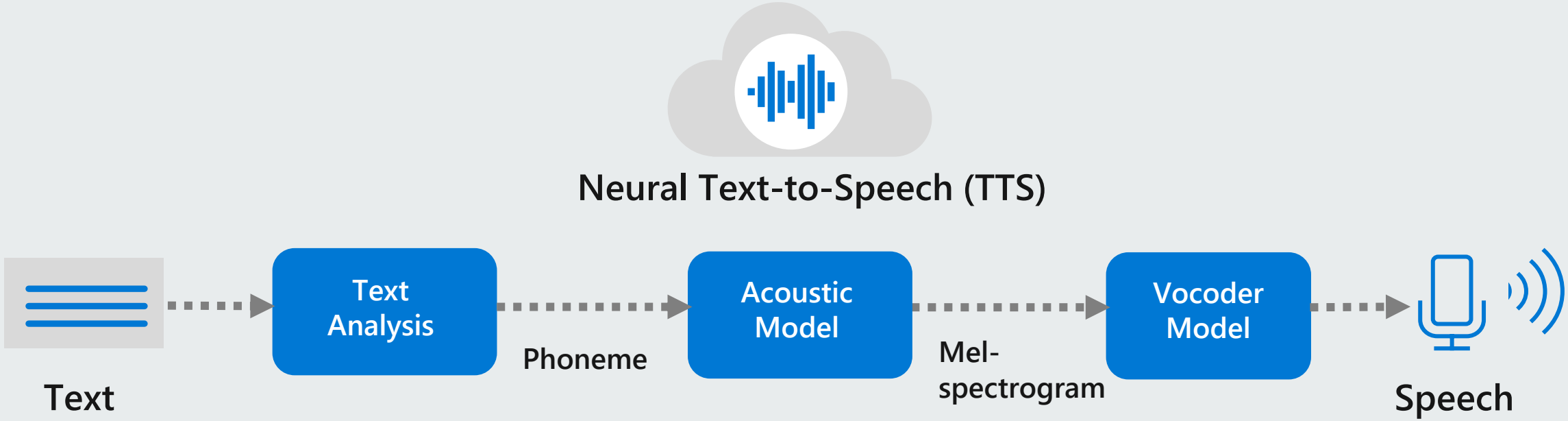
> Empowers students to access and engage with mathematical content and concepts in more personalized ways



# What is Neural Text-to-Speech?

Neural based end-to-end speech synthesis

- **Text Analysis**: text -> phoneme (e.g., Jan. -> January -> *daenjueri*
  - Text normalization, grapheme-to-phoneme conversion, polyphone disambiguation
- **Acoustic Model**: phoneme -> mel-spectrogram
  - Tacotron 2, DeepVoice 3, TransformerTTS, FastSpeech 1/2
- **Vocoder**: Mel-spectrogram -> waveform (neural model)
  - WaveNet, WaveRNN, LPCNet, WaveGlow, MelGAN



# Research related to “Immersive Reader” (1 of 2)

## Text-to-speech (Read aloud) and word or line highlighting

Increase reading speed and comprehension for all learners.

Support students with learning differences like Dyslexia.

Help emerging readers build confidence.

### 1 Font spacing

***10% faster and 50% fewer reading errors when reading text with optimized layout***

•Marco Zorzi, et al., Extra-large letter spacing improves reading in dyslexia

***Findings point to a subtype of dyslexia involving elevated crowding and demonstrate that individuals benefit from interventions personalized to their specific impairments.***

•Jason Yeatman, et. al., Optimizing text for an individual’s visual system: The contribution of visual crowding to reading difficulties

### 2 Line length

***27% increase in reading speed when using short line lengths***

•Matthew H. Schneps, et al., *Shorter Lines Facilitate Reading in Those Who Struggle*

### 3 Breaking words into syllables

***10% increase in reading comprehension for adults with syllabification***

•Yu-Chi Tai, et al., *Within-Word Text Segmentation on Lexical Processing and Reading*

***Building the syllabic bridge from available phonological syllables and frequent letter clusters may therefore be the first step in learning to read.***

•Nadège Doignon-Camus, et. al., The syllabic bridge: The first step in learning spelling-to-sound correspondences

### 4 Page color

***The children who chose a colored overlay read more slowly without the overlay than with it. These children reported more symptoms of visual discomfort and showed signs of tiring when they read without the overlay***

•Arnold J. Wilkins, et.al, Coloured overlays, visual discomfort, visual search and classroom reading

# Research related to “Immersive Reader” (2 of 2)

## Text-to-speech (Read aloud) and word or line highlighting

Increase reading speed and comprehension for all learners.

Support students with learning differences like Dyslexia.

Help emerging readers build confidence.

5 ***It was concluded that students with learning disabilities in all grades, except the 7th grade… benefitted from the read aloud accommodation more than their typical peers, presenting a differential boost.***

• Georgia Andreou, et.al., Accommodations on Reading Comprehension Assessment for Students with Learning Disabilities: A Review Study

6 **Simultaneous highlighting and voicing text –** “Dual highlighting is a related software feature, sometimes called masking, in which the context (sentence or paragraph) is highlighted in one color while the spoken word is highlighted in a second color, making it easier for readers to stay in sync with the spoken text….They found that students with LD who were given text passages with bimodal input performed as well on the comprehension questions as the average reader control group with visual input alone.

• *Heidi Pacuilla, Assistive Technology and Adult Literacy: Access and Benefits (2007)*

7 ***"Elkind examined the effectiveness of using speech synthesis during reading tasks on reading performance for post-secondary students with dyslexia. Their results showed participants not only demonstrated improved reading rates and comprehension, but also increased their ability to sustain attention while reading."***

• *Kim Floyd, et. al., The Efficacy of Assistive Technology on Reading Comprehension for Postsecondary Students with Learning Disabilities (2012)*

8 ***The use of screen readers or other text-to-speech software has led to increased reading comprehension performance for students with the weakest reading skills.***

• *Kim Floyd, et. al., The Efficacy of Assistive Technology on Reading Comprehension for Postsecondary Students with Learning Disabilities (2012)*

9 ***For students with reading disabilities, text-to-speech technologies may assist students with reading comprehension – a meta-analysis.***

• Sarah Wood, et. al., Does Use of Text-to-Speech and Related Read-Aloud Tools Improve Reading Comprehension for Students With Reading Disabilities? A Meta-Analysis

# Future Research – Neural Speech Synthesis

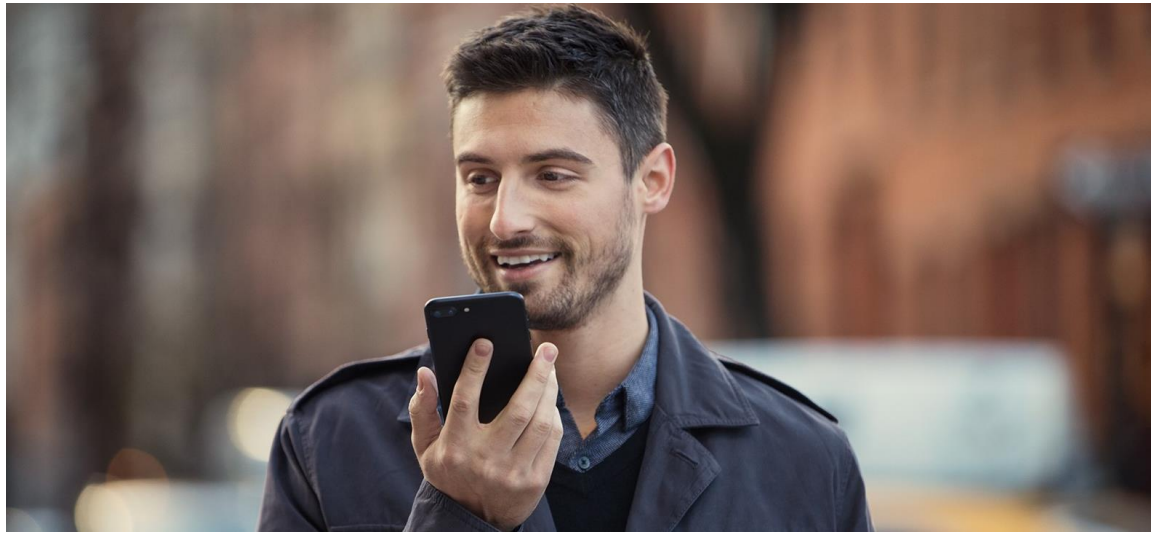
**A Survey on Neural Speech Synthesis, Xu Tan∗, Tao Qin, Frank Soong, Tie-Yan Liu, Microsoft Research Asia**
https://arxiv.org/pdf/2106.15561.pdf

Further research is required by both academic researchers and industry practitioners working on Text-to-Speech (TTS)

High-quality speech synthesis The most important goal of TTS is to synthesize high-quality speech. The quality of speech is determined by many aspects that influence the perception of speech, including intelligibility, naturalness, expressiveness, prosody, emotion, style, robustness, controllability, etc. While neural approaches have significantly improved the quality of synthesized speech, there is still large room to make further improvements.

- **Robust TTS**
- **Expressive TTS**
- **Parameter-efficient TTS**
- **Energy-efficient TTS**
- **Text-to-Video (talking heads)**
- **Responsible and Ethical Text-to-Speech**

# Inclusion and equity drives innovation for everyone



# Thank you