

A BRIEF GUIDE TO SPEECH RECOGNITION SOFTWARE IN TERTIARY EDUCATION

Lorraine Skelton

OTAGO POLYTECHNIC AUCKLAND INTERNATIONAL CAMPUS

Speech recognition software is rapidly expanding in the field of academia, becoming an essential tool for students, educators, and researchers alike. With the increasing demands on academic work—whether drafting research papers, transcribing interviews, or taking notes—efficient and accurate tools are effective in streamlining these tasks. Speech recognition software offers the ability to convert spoken words into text quickly and accurately, significantly enhancing productivity and facilitating many aspects of learning and research. This article discusses the ways in which speech recognition software can be used in academic settings and also reviews the best options currently available, evaluating each in terms of accuracy, ease of use, features, and cost-effectiveness. Popular options such as Dragon NaturallySpeaking developed by Nuance (<https://www.nuance.com>), Microsoft 365 Dictate (<https://www.office.com>), Google Docs Voice Typing (<https://docs.google.com>), Otter.ai (<https://www.otter.ai>), and Apple Dictation (<https://www.apple.com>) are reviewed to provide a guide for selecting the most suitable tool.

Note-Taking and Transcription: A study by Farrús et al (2023) notes that speech recognition software enhances note-taking, allowing students and academics to focus on the content rather than the act of writing or typing. The study also highlights the efficiency of automated transcription tools, emphasising their role in increasing productivity during lectures and meetings.

Accelerating the Writing Process: The software is beneficial for speeding up the writing process, particularly for those who find dictation a more natural form of expression. This method encourages a more dynamic approach to writing, making it easier to produce written content quickly. Research by Pennington and Rogerson-Revell (2018) discusses how speech recognition technology is increasingly integrated into academic writing processes, significantly reducing the time spent on drafting and editing documents.

Enhancing Accessibility: Accessibility in academia is greatly improved with speech recognition software, as it provides an alternative input method for individuals with disabilities. The study by Farrús (2023) highlights the inclusive benefits of speech recognition, especially for students with physical disabilities that make typing difficult.

Supporting Language Learning and Pronunciation: In language learning, speech recognition provides immediate feedback on pronunciation, aiding both educators and students in enhancing their spoken language skills. The review by Farrús (2023) explores the role of automatic speech recognition in language learning, focusing on its application in improving pronunciation and linguistic feedback, especially important for students who are studying in a second language.

Facilitating Research Data Collection: For qualitative research, speech recognition software automates the transcription of interviews and focus groups, which speeds up the data collection phase. The benefits of this automation are discussed in several academic works, including a study by Pennington and Rogerson-Revell (2018), which underscores the efficiency gains in research environments.

Boosting Productivity: By automating tasks such as writing and searching through documents, speech recognition software significantly enhances productivity. The integration of this technology into daily academic workflows is explored in research by Pennington and Rogerson-Revell (2018), which identifies productivity gains as a key advantage of using speech recognition in academia.

The next step is choosing the right software, from a variety of software options available, each with different features and pricing models. The choice includes decisions around accuracy, ease of use, compatibility with existing systems, and specific academic needs. There are several existing software products to consider, each with specific key criteria that may help in making an informed decision about which one to choose.

Dragon Naturally Speaking by Nuance is widely regarded as the industry leader in speech recognition software and has a lot of experience, having been developed in 1997. It boasts a high level of quality, especially around precision and customisation, offering up to 99% accuracy, with an ability to learn and adapt to the user's voice over time. Dragon also supports the creation of custom vocabularies and voice commands, which is particularly beneficial for users in specialised fields within the academic world, and having been purchased by Microsoft in 2021 it now has seamless integration with applications such as Microsoft Word and Google Docs, enhancing workflow efficiencies. It can also transcribe pre-recorded audio files, which is essential for researchers who conduct interviews or focus groups. The programme can also be used off-line. Dragon Naturally Speaking is however the most expensive option on the market at around NZ\$1,120, but it can be justified as an investment for those engaged in complex academic work, and is especially suited to researchers, academics, and students who require top-tier accuracy and customisable features for detailed projects.

Microsoft 365 Dictate is a built-in feature in Microsoft Office applications such as Word, Outlook and PowerPoint, making it easy to use for drafting emails, documents and presentations. It is a convenient and cost-effective option for users already subscribed to Microsoft 365, but this also means that you need to be online to use it. 365 Dictate offers real-time transcription with basic voice commands for punctuation and formatting, and supports multiple languages, catering to a diverse academic audience. Microsoft 365 Dictate is included in the Microsoft 365 subscription, which costs about NZ\$130 per year, which adds value for those already using this suite of programmes. Although it lacks the advanced features and accuracy of premium tools like Dragon, it is an accessible and efficient option for everyday academic tasks and can be suitable for students and educators who are already part of the Microsoft ecosystem and require a basic, integrated speech recognition tool.

Google Docs Voice Typing is a free tool available within Google Docs and provides a simple solution for users who need basic speech recognition capabilities without additional costs, making it an ideal option for budget-conscious users. It requires no additional setup or software installation and is easy to activate within Google Docs. It also benefits from Google's cloud-based collaboration features, enabling multiple users to work on a document simultaneously.

Google Docs Voice Typing offers the best cost benefit for users who need basic voice recognition without financial investment. While it lacks the advanced features and accuracy of paid software, its ease of use and integration with Google Docs make it a useful tool for students and educators looking for a free, straightforward speech recognition solution.

Otter.ai is a cloud-based transcription service known for its real-time transcription capabilities. It offers both free and premium plans, making it accessible to a wide range of users. It can be used for live transcription of lectures, meetings, and interviews in real-time. It also supports collaborative work on transcripts, which is particularly useful for group projects and can integrate with platforms such as Zoom for automatic transcription of virtual meetings. As Otter.ai offers a free tier with basic functionality and premium plans starting at around NZ\$160 per year, this makes it a cost-effective choice for those who need live transcription and collaboration features. Although it does not offer the same level of customisation as Dragon, its affordability and functionality make it an excellent option for students and academics, who need to transcribe live events and collaborate on notes.

Apple Dictation is a built-in tool available on macOS and iOS devices. It provides a convenient option for Apple users who need a basic dictation tool and works across all Apple devices, offering a consistent experience on MacBooks, iPhones, and iPads. Apple dictation is activated easily via a keyboard shortcut, making it suitable for quick dictation tasks and it supports offline dictation for short periods. Apple Dictation is free for all Apple users, making it an attractive option for those who need basic dictation capabilities without additional costs, but it is less suitable for longer, more complex tasks, as it lacks advanced features and higher accuracy.

CONCLUSION

Selecting the right speech recognition software will depend on individual need, the complexity of the tasks to be done, and budget constraints. Dragon Naturally Speaking remains the best investment, despite its higher cost. Google Docs Voice Typing stands out for its cost-benefit, offering a free and easy-to-use tool that meets basic academic needs. Otter.ai provides a good choice for those requiring live transcription and collaboration at an affordable price, and who need high accuracy and advanced features. Microsoft 365 Dictate and Apple Dictation offer cost-effective solutions within their respective ecosystems, making them suitable for users already utilising Microsoft Office or Apple products.

In summary, the best speech recognition tool is one that aligns with the specific demands of your academic work, offering the right balance between cost, accuracy, and functionality. By carefully considering these factors, academics and students can enhance their productivity and streamline their workflow.

REFERENCES

- Apple Inc. (2024). *Apple Dictation*. Retrieved from WWDC24 Highlights - Apple. <https://www.apple.com/newsroom/2024/06/wwdc24-highlights/>
- Farrús, M. (2023). Automatic speech recognition in L2 learning: A review based on PRISMA methodology. *Languages*, 8(4), 242. <https://doi.org/10.3390/languages8040242>
- Google LLC. (2024). *Google Docs Voice Typing*. Retrieved from VoiceTyper Dictation Software - Type with your Voice. <https://www.voicetyper.com/?src=ms&matchtype=p&kwd=voice%20typer&adgroup=Main%20Keyword%20Group&network=o&msclkid=ead037af76391afe554251578392a896>
- Microsoft Corporation. (2024). *Microsoft 365 Dictate*. Retrieved from Dictate in Microsoft 365 - Microsoft Support. <https://support.microsoft.com/en-us/office/dictate-in-microsoft-365-eab203e1-d030-43c1-84ef-999b0b9675fe>
- Nuance Communications, Inc. (2024). *Dragon NaturallySpeaking*. Retrieved from Dragon Speech Recognition - Get More Done by Voice | Nuance. <https://www.nuance.com/dragon.html>
- Otter.ai. (2024). *Otter.ai Live Transcription*. Retrieved from Otter.ai - AI Meeting Note Taker & Real-time AI Transcription. <https://otter.ai/?msockid=1d0fc4ef107a6a2434a2d03711ea6b6f>
- Pennington, M. C., & Rogerson-Revell, P. (2019). Using technology for pronunciation teaching, learning, and assessment. In *English pronunciation teaching and research*. Research and Practice in Applied Linguistics. Palgrave Macmillan. https://doi.org/10.1057/978-1-137-47677-7_5