Stop typing and work faster with speech recognition software

NUANCE

Introduction

Daily written office communication is both expensive and time-consuming. Faced with increasing communication requirements in a networked working environment, companies are scrutinising their existing workflows. In doing so, they are particularly focusing on three questions:

– How can companies reduce staff costs while increasing the productivity of their employees?
– Why should knowledge workers and managers be paid for time-consuming typing when they could be using their expertise in a far more productive way?
– How can high-quality documents be created more quickly and efficiently?

Digital speech recognition provides the answers to these questions. It enables emails and other documents to be created in a fraction of the time it takes to type. For example, in law firms and legal departments, many users are already using speech recognition software to create documents more quickly and to streamline workflows without changing existing business processes or IT systems.

This White Paper provides answers to questions about costs and productivity that companies need to ask. Find out about the advantages of speech recognition technology:

– **Create documents faster**
Convert the spoken word into written text with maximum recognition accuracy – up to 99 percent in almost all Windows® applications – and at triple the capturing speed.
– **Eliminate bottlenecks when creating documents**
The use of speech recognition reduces the time required for revising and correcting documents, accelerates completion and eliminates reliance on clerical staff.
– **Boost efficiency and profitability**
Greater efficiency and effectiveness in daily workflows through speech recognition provides the user with more time to focus on higher added value activities such as meetings with customers or product development.

How does speech recognition work?

Speech recognition software such as Dragon Professional Group by Nuance Communications uses the human voice as the main interface between the user and the computer. This user-friendly software uses highly developed technology to recognise and differentiate between the millions of human utterances that make up language.

On the basis of statistical models, the speech recognition program analyses an incoming acoustic signal, interprets this signal as a command or dictation, and converts it into digital information. This interpretation process is known as speech recognition. The success of the interpretation process is measured by the percentage of correctly recognised words.

Dragon is a speaker-dependent speech recognition system. It creates a speech profile for each system user. This individual speech profile contains information about the specific characteristics of the user's voice together with a number of user-specific words, identified as vocabulary. The speech profile also contains additional user-related information such as personal software settings and individually defined voice commands.

When users set and train their user profile, Dragon uses recognition models based on the analysis of thousands of voice and text samples. The created user profiles are then adapted according to their sound (acoustic model) and the words and expressions used (vocabulary and corresponding language model). This synchronising approach includes users' different accents and speech patterns.

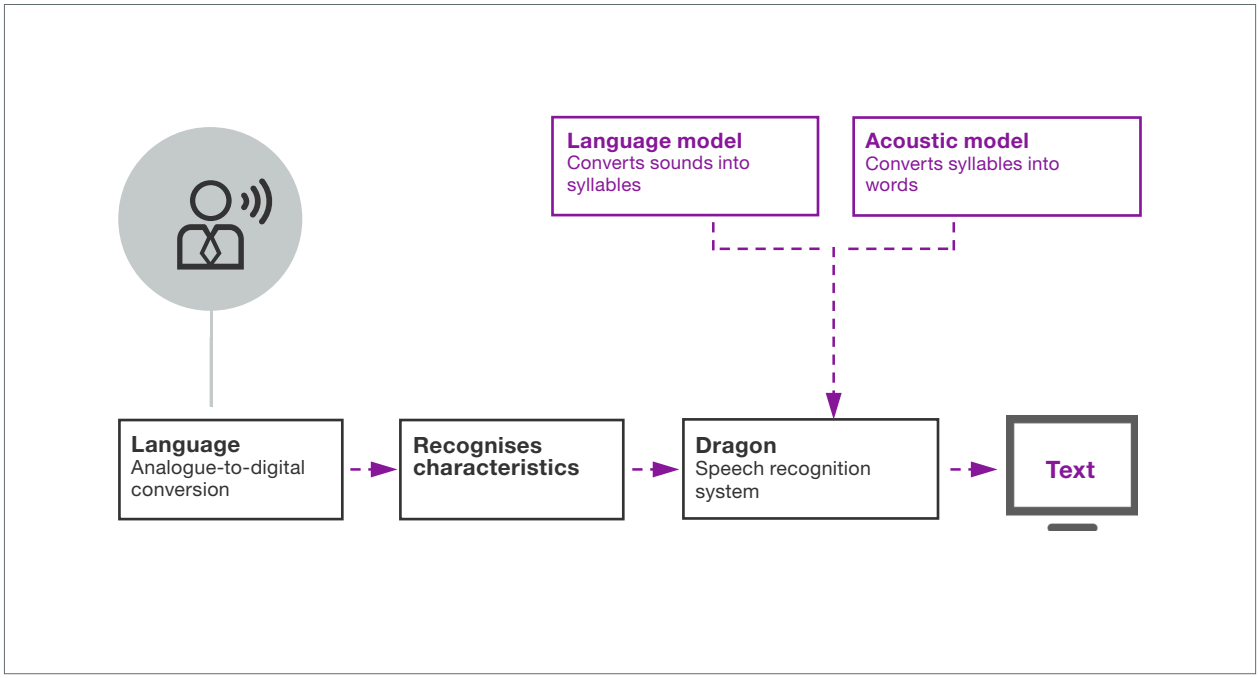


Fig. 1: Speech recognition operating mode.

During the recognition process, the software uses the user-specific voice profile to "guess" the spoken words. Each time a recognition error is corrected, the voice profile is updated, thus improving recognition accuracy (see Fig. 1).

Advances in the development of software and hardware have improved the recognition accuracy and user-friendliness of speech recognition technology. Therefore, it is well-worth taking another look at this technology, even if you had already made up your mind about speech recognition software in the past. The continuing developments of speech recognition mean that employees can create emails and other documents three times faster than by typing. Employees from all hierarchical levels and specialist fields can benefit from this technology. It is especially helpful to people with physical limitations.

What is speech recognition used for today?
- To create and edit all types of documents
- To manage emails
- To optimise third-party transcription
- To provide accessibility (support for users with disabilities, dyslexia etc.)
- To dictate information on the go and transcribe it later

How do business users benefit from speech recognition?

Create documents faster

Many managers and administrative staff members have to deal with a whole host of documents: memos, agreements, contracts, orders, reports, technical documentation, emails, etc. Companies take two different approaches to create all these texts:

– **Employees type out the information themselves**
  Not all employees are skilled at typing, so for many, manually entering texts is often a time-consuming affair. Furthermore, even employees who master the touch system cannot enter text as quickly as with speech recognition. Last but not least, typing can also be made difficult if the employee is suffering from an injury such as repetitive strain injury RSI – also known as mouse or tennis elbow.
– **Third-party transcription**
  The employee speaks into a Dictaphone. The audio file is transmitted to an external transcription office or in-house typist. This causes delays in the process as well as additional costs.

Dictating is three times faster than typing

Today, speech recognition systems are characterised by speed, accuracy and user-friendliness. They consequently offer a high-performance alternative to traditional document creation practices. In a nutshell, most people talk faster than they type.

The use of speech recognition enables employees to create electronic documents at a speed of up to 160 words per minute i.e. three times faster than typing. To do so, users simply talk into their computer, and the spoken words are immediately displayed in all Microsoft® Office® Suite applications, as well as Microsoft® Internet Explorer®, Corel® WordPerfect® and virtually all other Windows-based applications.

The use of a Bluetooth headset instead of a standard headset with a microphone provides an additional advantage: wireless dictation not only frees up employees' hands so that they can look over, for example, texts or notes that maybe lying on their desk. It also means that they can freely move about the office while dictating to fetch any additional references they might need.

Transcription: faster and better

Speech recognition is helpful for everyone who is regularly exposed to the pressures of the "Dictating-Transcribing-Revising-Correcting" production grind. Those who dictate their documents to a third party for transcription do not have to change their workflows. With speech recognition, the dictation is automatically transcribed. This simplifies and accelerates collaboration with office assistants who are responsible for transcribing and further processing the document.

The speech recognition system forwards the transcribed text file or synchronised audio file to the typist. Rather than having to transcribe the entire text from scratch, the typist opens the automatically generated transcription, listens to the associated audio replay while following the text on the screen, and makes any necessary corrections. Each correction performed improves the user's voice profile. In this way, the speech recognition system steadily increases recognition accuracy, ultimately saving even more time.

Using speech recognition software to enter text significantly reduces the processing time compared to traditional transcription processes. When transcriptions are completed in-house, the use of speech recognition technology reduces the workload and saves time for the typist who can focus on more productive tasks. Forwarding transcriptions for correction to third parties significantly reduces costs.

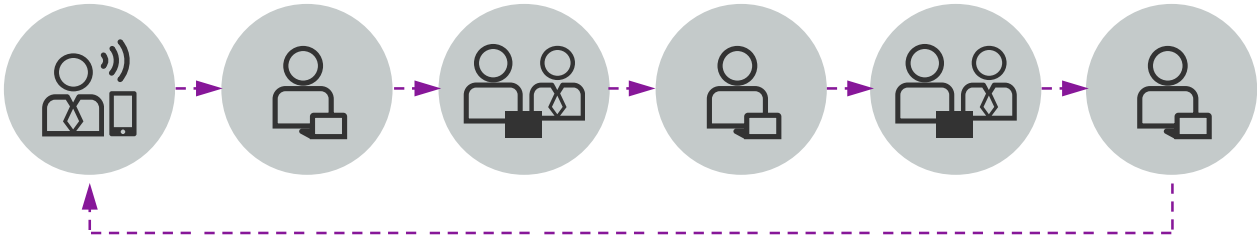
Dragon also offers further advantages with the "Auto-Transcribe Folder Agent". Using this function, the user can select a folder for incoming audio files. On receipt of an audio file, the system automatically converts it into a text file. The Auto-Transcribe Folder Agent enables the user to create an audio file from the converted text so that anyone wishing to make corrections can listen to the associated audio replay while reading the document.

With the aid of the "Dragon Voice Notations" function, users can make annotations which are saved as part of the audio dictation but do not appear in the written text. For example: "Please ask for Mr Smith's address by phone ..." the audio file is forwarded for further processing to the office assistants who import the recorded dictation using the creator's voice profile, convert it into text and finally make any corrections.

How accurate is speech recognition?

Speech recognition tools deliver up to 99% recognition accuracy right out of the box. The use of special vocabulary for specialist fields such as taxation or the insurance sector increases recognition accuracy. In addition, Nuance Partners work closely with companies right from the start to analyse in-house processes and identify approaches leading to increased productivity. As soon as such approaches have been established, customer-specific technical vocabulary is created and supplied together with customer-specific commands and macros. This leads to a significant increase in productivity, especially when used across the company.

Using traditional transcription



1. The employee dictates a memo consisting of three paragraphs.
2. The recording is forwarded to the clerical staff.
3. The clerical staff listen to the recording and type out the text.
4. The clerical staff deliver the draft to the author after approximately 1 to 2 hours, depending on the volume of work.
5. The employee interrupts his current activity, reads the memo and makes any changes.
6. The clerical staff process the changes, receive any explanations from the employee and complete the document within 1 to 2 hours, depending on the volume of work.
7. The clerical staff forward the revised memo to the employee for review.
8. The employee interrupts his current activity, reads the document and confirms any amendments made.
9. The clerical staff print out and send off the finished memo.

Total time taken: 2-6 hours, depending on the volume of work.

Using speech recognition

1. The employee dictates a memo consisting of three paragraphs into a digital recording device or into a microphone connected to the PC.
2. During synchronisation with the PC, Dragon imports the recording from the recording device for automatic transcription.
3. The clerical staff revise the text with the help of the associated audio file within a period of 15-60 minutes, depending on the volume of work.
4. The clerical staff forward the created document to the employee for review.
5. The employee interrupts his current activity, reads the memo and makes any changes.
6. The clerical staff print out and send off the finished memo.

Total time taken: 0.5-1.5 hours, depending on the volume of work.

Customising vocabulary and technical vocabulary

Every industry uses specific names, technical terms and abbreviations based on their area of specialisation. These specific expressions are frequently used in legal documents, letters and other documents. Dragon learns from you. Once you dictate these expressions, you can correct them if needed, thereby adding them to the system. This process of gradual adaptation can be significantly reduced and optimised by using targeted customised vocabulary. To do this, existing documents are checked, grammatically analysed and processed into individual vocabulary. Nuance Partners offer this service and make technical vocabulary from different sectors available as a basis for customisation, therefore significantly reducing expenses. Vocabulary adapted to your specific industry can reduce errors by up to 30 percent.

Automating workflows

Speech recognition software can do much more than just converting dictations into text. Solutions such as Dragon Professional Group help users deal with a multitude of routine communication tasks more quickly and efficiently.

Commanding applications by voice

Use simple voice commands to create files, send emails, schedule tasks and meetings, open and close applications, save and convert documents and perform searches on the Internet or your computer with remarkable speed and ease. Many applications can become more user-friendly and effective when used in combination with speech recognition technology. Full text searches, database queries and filling out forms can be performed faster by voice than by typing. Similarly, document administration, document creation and its automation are particularly well-suited to being used and managed through voice commands.

Routine tasks such as text and data entry can be performed through voice commands in most programs without any adaptation. Additional functions of different programs can be executed using macros which are run by voice commands.

Managing email

The daily cost of managing electronic mail is constantly increasing. This can be counteracted with the use of speech recognition software for writing, sending and answering emails, and navigating in common email programs such as Microsoft® Outlook®. In addition, certain speech recognition programs include the “read back function”, whereby users have the opportunity to have their emails read aloud, so that they can deal with other tasks while listening.

Working with speech on the Internet

It is easier than ever before to use speech recognition programs to carry out searches on the Internet, collect information and browse websites. They are not only limited to browsing on the public Internet, but can also be used for private Intranets and other HTML interfaces.

Bringing together multi-step tasks into one simple voice command

Dragon command shortcuts

Most speech recognition programs enable users to speak out a standard voice command which instructs the computer to perform an action. For example, a user says “Start WordPerfect”, upon which the PC runs this program. Technologically advanced speech recognition programs such as Dragon Professional Group also enable users to condense several work steps into individual voice commands. With the integrated script tool (Microsoft® VBA-compatible), you can program recurring keyboard or mouse inputs and perform them with a single voice command. In this way, any application or form can be browsed and controlled by voice commands.

For example, an employee uses Dragon command shortcuts to process emails, simply by saying: “Send email to John Smith and Raphael Sanchez” or “Search maps for financial advisers in London” and Dragon makes it happen. He can also say “Open top site for Nuance” and Dragon automatically opens the top-ranked page for your keyword. In addition, he can easily create or import custom voice commands that let him fill out forms, insert frequently used text and graphics, control application elements such as menus and toolbars or automate business processes, meaning that, with a single voice command, a document can be completely created and individually adapted by browsing different input fields and entering different information.

Boilerplate text as voice commands
Many documents created by companies always contain the same recurring, standard elements. That means that many employees enter the same information again and again into their documents. To counteract this and to complete documents more quickly and efficiently, users can create boilerplates and add these to documents or emails by means of voice commands. For example, the boilerplates contain regularly used expressions and formulations or even graphics. These user-specific commands can also contain "Dragon templates".

Macro commands
Entering data or filling out forms are typically repetitive work steps which can be accelerated with the use of speech recognition. When users find sophisticated software programs difficult to use, they may find it easier to "talk" to the computer rather than controlling the program by keyboard or mouse. The creation of macros enables users to move smoothly between input fields or to perform a sequence of keyboard inputs or mouse movements by simply using their voice. Even experienced typists are often slower when typing numbers and letters. In addition, spreadsheets can be created and processed by voice. Accounting and time keeping programs can also be controlled by voice.

With the aid of Microsoft® Visual Basic®, experts and secretaries – or more precisely the IT departments that support them – create, for example, a voice command which saves a document, sends it to a standardised mailing list and prints it out at the head office and in a specific branch. All the steps of this workflow are performed with a single spoken command such as "Complete document". If you require support when developing speech macros, Nuance Partners can support you directly or provide training.

Increasing productivity on the go

Speech recognition technology can also increase productivity during business trips or when carrying out activities outside the office. The business traveller dictates his text into a digital Dictaphone or other recording device, so that it can be automatically transcribed on the PC later. As Dragon does not just transcribe but also records the dictation as an audio file, it can also be used by third parties to correct the transcription. When speech recognition is used on a company-wide level, the speech profiles can be accessed anywhere and at any time on any PC or laptop in the company network. Furthermore, the integration of Dragon into different dictation management systems simplifies collaboration with external clerical staff. Nuance Partners will be happy to assist you in your selection of the most suitable system along with the certified dictation hardware.

Server profiles for using Dragon Professional Group on several PCs

Dragon Professional Group can be used by one user with only one Dragon licence on several PCs (for example on an office PC, a home PC, a laptop when on the move and also on the PC of the member of clerical staff making any corrections) as long as the software is used by just one speaker. Using the server profile functionality and synchronising a central profile, Dragon can ensure that adjustments and corrections to improve speech recognition quality are available not only on the PC on which they were effected, but also on all of the user's other PCs. By installing Dragon on a Citrix server computer, you can have Dragon rely on the Citrix server for all processing activities.

Managing a desktop solution

When speech recognition is successfully used in the workplace, training and adapting the system to customer-specific requirements play an important role. Some companies decide to manage the installation, company-specific adaptation and training of their speech recognition systems themselves. However, most companies prefer to ask a Nuance Partner for advice in these areas.

In general, an evaluation of the necessary system requirements is recommended in order to use the speech recognition software. Operating speech recognition software requires a high level of processing power. To summarise, the faster the processor, the better the recognition performance. Furthermore, users wishing to work with several applications at the same time benefit from increased memory (RAM).

Managing networks
Using a PC dictation solution from a central network location for installation and management, system administrators can:

– create and manage user profiles in the network
– share user-specific vocabulary and commands automatically
– manage settings
– restrict certain functions to certain roles
– automatically synchronise updates and changes via a number of communication protocols
– perform system backups.

The latest speech recognition systems include administration tools enabling all users in the company to be provided with user-specific vocabulary and macros. For example, an administrator can transfer improvements (which one user has made to the vocabulary and saved in his user profile) to another user. In this way, a number of different users can be provided with updates to shared vocabulary. This eliminates time-consuming input of individual new words and pronunciations which each individual user enters into the PC himself.

In addition, the central network administration enables users to share tried and trusted methods with other users quickly and easily via the network.

User expectations and training

Most speech recognition systems win over new customers with high-performance functions for dictation and application controls – right from the very first time they are used. However, initial professional training and customisation can help companies significantly increase productivity and save money, which ultimately leads to a return on investment (ROI). Initial training brings rapid success, helps the user feel more self-confident, reduces support and maintenance costs, promotes the success of pilot projects and secures investments. Regular on-going training and advanced training modules enable users to become familiar with enhanced software features and functionalities and to enhance their productivity even more.

In addition, defining realistic user expectations directly influences the success or failure of the speech recognition program. Through the use of speech recognition systems, companies can automate their business processes without interrupting existing routines. However, when doing so, they should keep in mind that the introduction of new technology requires a transitional phase, including familiarising employees with the new “Dictate don’t type” approach.

At this point, professional training can facilitate change management. Training accelerates and simplifies the company’s entire transitional phase, as well as increases users’ acceptance. Most users who are familiar with dictating into an electronic device quickly get used to using speech recognition software with little difficulty. However, they may have fallen into the habit of mumbling and slurring their speech when dictating, assuming that any inaccurate pronunciation will be correctly interpreted during the transcription. With speech recognition, the quality of the acoustic speech signal is just as important as the quality of the sound card.

Customisation
Investing in specific vocabulary can also make an impact. By customising technical or thematic system vocabularies before using the speech recognition software, companies can achieve remarkably accurate recognition performance even when using the system for the first time. Nuance Partners work closely with companies right from the start to analyse in-house processes and identify approaches leading to increased productivity. As soon as such approaches have been established, customer-specific technical vocabulary is created and supplied together with customised commands and macros that considerably accelerate the execution of routine multi-level tasks.

The resulting saved time then leads to a considerable increase in productivity if the system is used on a company-wide level. Text macros and step-by-step macros can be created without prior knowledge. In contrast, sophisticated macros require advanced scripting knowledge when using Microsoft® Visual Basic®. Certain companies have IT departments which are able to create sophisticated macros for Dragon users after initial training. Alternatively, companies can ask a Nuance Partner to create one or all initial commands and to train a few users as experts.

Does the expenditure pay off?

Using a speech recognition system on a company-wide level consists of the following components:

– custom-specific software
– professional services for planning, installation, product adaptation, training and technical support
– audio hardware such as microphones, digital recording devices, wireless microphones
– IT resources such as server and memory systems, back-end and system integration, data and user profile administration, user support

An assessment of the return on investment (ROI) should include costs of the staff members required to perform the transcription work. In addition, further staff-related costs must be taken into account:

– Costs due to repetitive strain injury (RSI) and similar symptoms caused by working on a computer
– Estimated annual costs for loss of employee productivity due to RSI

In the USA, but also in other countries throughout the world, RSI is by far the most common work-related medical condition and cause of inability to work. Many employees suffering from RSI are not able to operate a PC with their hands. Speech recognition can counter the consequences of RSI by preventing illnesses and serious health problems from occurring in the first place. The use of speech recognition helps injured employees return more quickly to their employment, therefore reducing workers' compensation insurance and legal health costs for employers and staff expenses.

About Nuance Communications

Nuance is the leading provider of voice and imaging solutions for businesses and consumers around the world. Our technologies, applications and services make the user experience more compelling by transforming the way in which people interact with information and how they create, share and use documents. Every day, thousands of users and hundreds of businesses benefit from Nuance's proven, high-performance applications and customer services.

For more information, please visit **www.nuance.co.uk**.

About Nuance Communications, Inc.
Nuance Communications is reinventing the relationship between people and technology. Through its voice and language offerings, the company is creating a more human conversation with the many devices, electronics, apps and services around us. Every day, millions of people and thousands of businesses experience Nuance through intelligent systems that can listen, understand, learn and adapt to your life and your work. For more information, please visit nuance.co.uk.

NUANCE