IBM VoiceType Dictation for Windows 1.32 Recognizes Speech Quickly, Accurately

VoiceType™ Dictation — the highly accurate speech recognition solution — now runs on Windows† 3.1x and Windows 95[1]. VoiceType Dictation provides a major improvement in how you work with your computer. With VoiceType Dictation, you can talk directly into your computer and create text faster than many people can type. As you talk, text is created in a Dictation Window that supports a rich set of document editing and formatting functions. The text can be saved to a file or integrated with other applications.

This release supports many "hands-free" enhancements, such as the ability to turn the microphone on and off by voice and to navigate through the application's menus by speaking "discrete" numbers and letters (with distinct pauses between words). You can also control the Windows Program Manager in Windows 3.1x and the Task Bar and Desk Top in Windows 95 and applications by speaking U.S. English (or Spanish with the Spanish version).

When running under Windows 95, VoiceType Dictation has the features of a full Windows 95 application, including:

- Long filenames
- Object linking and embedding (OLE) 2.0 container and object support, and OLE 2.0 drag-and-drop support
- E-mail (MAPI) supported directly from the Dictation Window
- Auto-install from a CD-ROM

How Does It Work?

A sophisticated statistical language model that understands common word pattern usage and a detailed understanding of your voice allows VoiceType Dictation to accurately convert your speech to text. Words that sound alike such as *to, too,* and *two* are easily handled.

The product includes a general office vocabulary of over 22,000 words. You can add up to 2,000 more words while dictating. To make the product even more accurate in professions like medicine, journalism, and law, you can expand the product with specialty vocabularies.

With VoiceType Dictation, there is no need to stop and make corrections as you are dictating; you can edit after you complete your thoughts. VoiceType can even replay what you said, so that you never have to rely on memory to get the details right.

One-Time Charge: $599 (Software only)

Description	Planned Availability

VoiceType Dictation for Windows 1.32
 - Software only, Software only Upgrade,
   ISA Solution Pac[2], MCA Solution Pac[2],
   or PCMCIA Solution Pac[2]
	November 30, 1995

These products will be available in Spanish on March 29, 1996.

™   Trademark of International Business Machines Corporation
†   Trademark of Microsoft Corporation
[1]   Special bid requests will be accepted for support of the Windows NT operating system.
[2]   Includes software, VoiceType microphone, and VoiceType Dictation Adapter

---

IN BRIEF . . .

This high-accuracy speech recognition solution runs on Windows 3.1 and Windows 95, using proven technology based on more than two decades of IBM experience. With VoiceType Dictation, you can:

◆ Talk directly into your computer, creating text that is accurate and available within seconds

◆ Capture precisely what you say because sophisticated technology accurately converts your speech to text

◆ Navigate in standard Windows applications by speaking the words that appear in the menu bar, command buttons, check boxes, and option buttons

◆ Improve recognition accuracy in certain professions by adding specialty vocabularies

◆ Maintain your train of thought and create text while busy

Operating Environment: PC

Description

VoiceType Dictation for Windows is a highly accurate, large-vocabulary speech recognition program that converts speech to text. In many cases it can recognize words that are run together, slurred, or partially obscured by background noise. The product includes a language model that understands common word pattern usage.

VoiceType Dictation includes a general office vocabulary. IBM has optimized the size of the vocabulary for performance and accuracy. This optimization in combination with the language model results in a program that can be faster and more accurate than products with a larger vocabulary. You can also add up to 2,000 of your own words to the vocabulary.

Dictation ─ General Office Users

The general office vocabulary can be utilized by a wide variety of users who need to create and edit text. Users who would be especially interested are corporate users of e-mail and word processing systems, executives, professionals, students, technical writers, and translators. The ability to use verbal short cuts for distribution lists, standard headings, and command sequences can increase the productivity of these users. In addition, users who type slowly can achieve greater levels of speed and accuracy.

The sophisticated, adaptive capability of the system to learn users' dictation styles and accents can allow computer users who are speaking in secondary languages to achieve greater levels of speed and spelling accuracy. Technical writers can easily add special and foreign words to the system.

The dictation function includes a 22,000+ word general office vocabulary supporting U.S. English or Spanish dictation and a spelling dictionary of 100,000 words. Text is dictated into the Dictation Window and can be transferred into applications running on Windows.

VoiceType Dictation temporarily stores your spoken words. You can listen to your spoken words to assist you in correcting your dictated text. A loudspeaker is required when using this feature if your microphone does not contain an earpiece for feedback. Because the spoken word file can be quite large, the spoken word acoustic file is erased once the dictation is corrected and stored as coded text. The coded text file is much smaller than the spoken acoustic file, using far less disk storage.

Increased Productivity

General Office

- Eliminates or reduces time spent proofreading, correcting, typing, and waiting for transcribing.

- Captures information faster from individuals who have a limited ability or reluctance to use keyboards.

- Potentially improves end-user quality and information content of correspondence, reports, and similar documents.

- Enables greater speed and spelling accuracy for users who are working in secondary languages.

Medical

- Immediate availability of documents reduces the amount of transcription proofreading and allows immediate transmission to co-workers.

- Dictation can be performed while wearing gloves or working on a patient without looking at the screen or touching the keyboard or mouse.

- The confidentiality of patient notes and charts may be increased because of reduced reliance on outside transcription services.

- Costs can be reduced for transcriptions.

Legal

- Opportunity for improved productivity and reduced transcription costs.

Journalism

- Opportunity for improved productivity.
- Reduced reliance upon the keyboard.
- Significant number of proper names and places reduces spelling errors.

Voice Manager: The Voice Manager is a key part of VoiceType Dictation. The Voice Manager provides two broad categories of speech-related capabilities:

- Navigation
- Dictation

The Voice Manager Navigation function automatically speech-enables the Windows Program Manager and Windows applications that internally use the standard Windows user interface controls, such as menus, dialog boxes, and list boxes. Users can, for example, speak the words that appear in the menu bar, command buttons, check boxes, and option buttons of standard Windows applications. Speech commands are also available to scroll client application areas and list boxes.

To use navigation functions, the national language of the operating system and the selected Windows applications must correspond to the language version of VoiceType Dictation for Windows 1.32 unless the user trained the program to recognize the specific words.

Other Windows applications can be speech-enabled by defining voice macros for them. Voice macros allow users to define a sequence of actions that can be subsequently initiated with a simple spoken command. The macros can also emulate keyboard functions.

The Voice Manager provides function for the creation of text macros and for the easy transfer of dictated text to other applications. Text macros allow users to create standard text that can be included in the Dictation Window with a simple spoken command. Use of text macros can significantly increase the user's overall dictation speed. Dictated text can be transferred to another application such as a word processor through the use of drag-and-drop or cut-and-paste.

Speech-Enabled Applications: The Voice Manager provides a quick and easy way for users to bring the power of speech recognition to existing applications. The Voice Manager supports speech-enabling applications through three of its functions:

- Automatic enablement
- Command macros
- Text transfer

An application that responds to voice through the use of these functions is speech-enabled.

The U.S. English version of VoiceType Dictation for Windows has been tested with the following applications. The listed applications are compatible with the Voice

Manager functions of automatic enablement, command macros, and text transfer.

- Word Processors

    - Microsoft** Word** for Windows 6.0 and 7.0
    - WordPerfect** for Windows 6.1
    - Lotus** AmiPro** for Windows 3.1
    - Lotus WordPro** 96

- Productivity Applications

    - Lotus Notes** 3.0
    - Lotus cc:Mail** 2.2
    - Microsoft Works** for Windows 3.0

- Spreadsheets

    - Microsoft Excel** for Windows 5.0 and 7.0
    - Lotus 1-2-3** for Windows 5.0

- Graphics

    - Microsoft Powerpoint** for Windows 4.0
    - Lotus Freelance Graphics** for Windows 2.1
    - CorelDraw 5.0 and 6.0

- Desktop Publishing

    - Aldus Pagemaker 5.0

For navigation, the national language of the operating system and the application must correspond to the language of VoiceType Dictation for Windows 1.32 unless the user trained the program to recognize these specific words.

Additional Windows applications may work, though IBM has not tested them at this time.

Investment Protection

With VoiceType Dictation for Windows, existing applications may be protected and extended. Text created by voice can be passed directly to applications for document production without an intermediate step such as transcribing.

Product Positioning

IBM now offers the following speech recognition dictation products:

- VoiceType Dictation for Windows
- VoiceType Dictation for OS/2®

These products require discrete speech (distinct pauses between words) and support recognition of very large vocabularies and speech navigation.

The VoiceType Dictation for Windows and VoiceType Dictation for OS/2 (OS/2 2.1 and Warp) are stand-alone products with the same function. Selection should be made based on your operating system environment. VoiceType Dictation for Windows Version 1.32 replaces VoiceType Dictation for Windows (Version 1.1).

IBM, in addition, offers the speech recognition product VoiceType Application Factory (VTAF) for Windows.

IBM VTAF for Windows lets software developers add speech recognition to the software they develop. VTAF includes a sample speech-aware desktop navigator application that allows you to command and control Windows 3.1 applications using voice commands. VTAF operates with a well-defined grammar and highly structured, predictable vocabulary, while the VoiceType Dictation for Windows 1.32 and VoiceType Dictation for OS/2 1.1 are dictation products with a large flexible

vocabulary for general office use. In addition, industry-specific language models are available for VoiceType Dictation for Windows 1.32 and VoiceType Dictation for OS/2 1.1. VTAF is not an end-user product but is intended for developers to imbed English speech recognition function into their applications.

Publications

The following publications will be available from IBM. To order, contact your IBM representative or call 800-879-2755.

Title	Order Number
IBM VoiceType Dictation Fact Sheet	G221-3915
IBM VoiceType Dictation for Windows 1.32 User's Guide	S83G-7939
IBM VoiceType Dictation for Windows 1.32 Quick Reference Card	S52H-3625

One copy of the following publications are supplied automatically with the basic machine-readable material of VoiceType Dictation for Windows 1.32.

IBM VoiceType Dictation for Windows 1.32

	Order Numbers	
Language	User's Guide	Quick Reference Card
Spanish	SH10-9106	SH10-7974
U.S. English	S83G-7939	S52H-3625

Education Support

Call IBM Education and Training at 800-IBM-TEACH (426-8322) for education catalogs, schedules, and enrollments.

Technical Information

Specified Operating Environment

Hardware Requirements

Minimum Configuration

- A 486SX-25MHz PS/2®, ValuePoint™ system, Aptiva™, PS/1®, ThinkPad®, or IBM compatible

- At least 8MB of available memory (above operating system and other application requirements)

Recommended Configuration

- A 486DX-33MHz PS/2, ValuePoint system, Aptiva, ThinkPad, or IBM compatible

- At least 12MB of available memory (above operating system and other application requirements)

All Configurations also Require

- At least 35MB to 45MB of available disk space (plus an additional 42MB to 45MB during the enrollment process), depending on the language chosen

- When running Windows 3.1x, the Windows disk space allocated to permanent virtual memory should be set to at least 16MB. Permanent virtual memory is disk space used by the operating system as if it where

actually memory. Invoke the Windows control panel to set the size of the disk allocation.

- An appropriate IBM VoiceType Dictation Adapter (which is included in the Solution Pacs)

  Note: Personal Dictation Adapters are compatible with VoiceType Dictation 1.xx.

- An available ISA slot for the VoiceType ISA Adapter, or an available Micro Channel® slot for the VoiceType Micro Channel Adapter, or an available Type II PCMCIA slot for the VoiceType PCMCIA Adapter

- A microphone (which is included in the Solution Pacs)

- A loudspeaker (optional but recommended) and cables compatible with the audio subsystem (if the microphone used does not contain an earpiece for feedback)

Software Requirements

- Microsoft Windows 95 is a prerequisite when using the VoiceType ISA Adapter, the VoiceType Micro Channel Adapter, or the VoiceType PCMCIA Adapter.

  or

- IBM DOS 5.02 or MS DOS 5.0 and Windows 3.1/3.11 or Windows for Workgroups 3.1/3.11 is a prerequisite when using the IBM VoiceType ISA Adapter or the IBM VoiceType Micro Channel Adapter.

  or

- IBM DOS 6.1 or MS DOS 5.0, or later, and Windows 3.1/3.11 or Windows for Workgroups 3.1/3.11 is a prerequisite when using the VoiceType PCMCIA Adapter. The VoiceType PCMCIA Adapter requires card and socket services, such as those included with DOS 6.1, that meet the Personal Computer Memory Card International Association 2.0x standard. Additionally, PCMCIA device drivers may be required from the manufacturer when using non-IBM systems.

VoiceType Dictation for Windows 1.32, U.S. English is a prerequisite for the optional VoiceType Vocabularies for Dictation 1.1. Refer to Software Announcement 294-706, dated November 8, 1994, for information.

To use navigation functions, the national language of the operating system must correspond to the language of VoiceType Dictation for Windows 1.32 unless the user has trained the program to recognize the specific words.

Compatibility: The U.S. English version of VoiceType Dictation for Windows 1.32 has been tested and found compatible with the following applications:

- Word Processors

  - Microsoft Word for Windows 6.0 and 7.0
  - WordPerfect for Windows 6.1
  - Lotus AmiPro for Windows 3.1
  - Lotus WordPro 96

- Productivity Applications

  - Lotus Notes 3.0
  - Lotus cc:Mail 2.2
  - Microsoft Works for Windows 3.0

- Spreadsheets

  - Microsoft Excel for Windows 5.0 and 7.0
  - Lotus 1-2-3 for Windows 5.0

- Graphics

  - Microsoft Powerpoint for Windows 4.0
  - Lotus Freelance Graphics for Windows 2.1
  - CorelDraw 5.0 and 6.0

- Desktop Publishing

  - Aldus Pagemaker 5.0

Refer to Hardware Announcement 194-394, dated November 8, 1994, for systems tested.

The IBM Personal Dictation System U.S. English vocabularies are not compatible with VoiceType Dictation for Windows 1.32 U.S. English.

The IBM Personal Dictation Adapters are compatible with VoiceType Dictation for Windows.

Limitations: Limitations with VoiceType Dictation for Windows 1.32:

- If another language should be installed together with U.S. English, U.S. English must be installed first. Languages other than U.S. English can be installed in any order.

- During the training and enrollment process, VoiceType Dictation uses a DOS memory extension program included with the product. Other memory extension programs should not be active during the training and enrollment process.

- VoiceType Dictation and its files should not be compressed with disk compression utilities.

- VoiceType Dictation may not work with all processor boards.

- VoiceType Dictation for Windows 1.32 is not compatible with WIN-OS/2™.

Performance Considerations: Speech recognition requires significant system resources to provide a satisfactory response. Particular attention should be paid to the system requirements, which should be considered minimum system requirements. Performance can be influenced by many factors, but three key areas are processor power, available memory, and virtual memory space.

As for processor speed, faster machines provide a quicker response to any user-initiated function whether it be by keyboard or spoken commands. As with a keyboard, processor speed does not affect recognition accuracy. On a faster machine, you will notice all speech recognition functions completed faster. For example, if you speak at a specific dictation rate, then on a faster machine, translation of the spoken words into recognized text will complete faster. For rapid input on a slow processor, you can notice that there is a delay after the dictated text finally appears on the screen. The delay is less on faster processors.

If you use relatively large programs (such as a word processor) with the speech recognition system, then sufficient RAM for both programs is useful to prevent excessive swapping. (You can notice the effect of swapping even on word processors alone if real RAM is reduced.) In addition, sufficient contiguous disk space and a 32-bit file system greatly improve the ability to do swapping rapidly.

Refer to the **Technical Information** section for the recommended hardware configuration.

Planning Information

Customer Responsibilities: The customer must provide at least the minimum hardware and software environments in which the VoiceType Dictation for Windows 1.32 will operate.

Each user must enroll on the system to insure the system can recognize his or her pronunciation. The enrollment process also helps users learn the "discrete" speech pattern understood by VoiceType Dictation, where users leave a short pause between words. This results in very high accuracy and near-instant recognition.

If the microphone used does not contain an earpiece for feedback, the customer is responsible for selecting and installing an optional loudspeaker and the related cables. The loudspeaker does not need to be expensive, but it should have at a minimum the following characteristics:

- Amplified loudspeaker
- Monaural (stereo speakers are not required)
- Volume and tone controls

An example of a loudspeaker that has performed satisfactorily is the Radio Shack** 32-2040 compact one-watt speaker/amplifier or Sony** SRS-58PC.

IBM recommends that the customer use the microphone in the Solution Pac from IBM. If the customer prefers a different microphone, care should be used since the choice of a microphone can influence speech recognition accuracy. An alternative microphone should be very high quality and designed for close-range pickup to reduce background noise. The microphone must have a flat audio frequency response range. It is recommended that enrollment be performed with the alternative microphone selected. One sample of microphone characteristics that have been tested with VoiceType Dictation for Windows 1.32, U.S. English is given below:

- Type: dynamic

- Frequency response: 50 to 15,000 Hz

- Polar pattern: cardioid (unidirectional) response, uniform with frequency, symmetrical about axis

- Impedance: 150 ohms for connection to microphone inputs rated at 19 to 300 ohms

-

> Output level: (1000 Hz)

> Open Circuit voltage -47.0 dB (4.5 mV)

> Power level -66 dB

- Sensitivity: (Hum) 38.4 dB equivalent SPL in a 1 millioersted field

Packaging: VoiceType Dictation for Windows 1.32 Software only and VoiceType Dictation for Windows 1.32 Software Upgrade is distributed in one package with the following materials:

- VoiceType Dictation for Windows 1.32 User's Guide
- License Information
- Proof of License or Proof of Entitlement
- CD-ROM
- 3.5-inch (1.44MB) installation diskettes
- Dictation Quick Reference card
- Registration card
- CompuServe** Information Service Flyer (U.S. English only)

VoiceType Dictation for Windows 1.32 Solution Pacs are distributed in one package with the following materials:

- VoiceType Dictation for Windows 1.32 User's Guide
- License information
- Proof of License or Proof of Entitlement
- CD-ROM
- 3.5-inch (1.44MB) installation diskettes
- VoiceType qualified microphone with cable and warranty
- Dictation Quick Reference card
- Registration card
- CompuServe Information Service Flyer (U.S. English only)
- VoiceType Dictation Adapter (ISA or MCA or PCMCIA)

Security, Auditability, and Control

The announced programs use the security and auditability features of the operating system and IBM system unit in which they are installed.

User management is responsible for evaluation, selection, and implementation of security features, administrative procedures, and appropriate controls in application systems and communication facilities.

Ordering Information

Orders for new licenses will be accepted now.

Shipments will begin on the planned availability date.

To order these programs for order type, specify the order type number, feature number, part number, program name, media, and quantity.

Program Package: A 3.5-inch program diskette and CD-ROM with documentation.

Description	Order Type	Part Number
VoiceType Dictation for Windows 1.32, U.S. English		
Software only	5801-AAR	52H3636
Software only Upgrade	5803-AAR	52H3640
ISA Solution Pac	5801-AAR	52H3630
MCA Solution Pac	5801-AAR	52H3632
PCMCIA Solution Pac	5801-AAR	52H3634
VoiceType Dictation for Windows 1.32, Spanish		
Software only	5801-AAR	1957791
Software only Upgrade	5803-AAR	1957862
ISA Solution Pac	5801-AAR	1957855
MCA Solution Pac	5801-AAR	1957858
PCMCIA Solution Pac	5801-AAR	1957860

Documentation: Separate documentation with no media.

Description	Order Type	Part Number
VoiceType Dictation for Windows 1.32		
User Guide, U.S. English	5806-AAR	52H3623
Quick Reference Card, U.S. English	5806-AAR	52H3625
User Guide, Spanish	5806-AAR	1957792
Quick Reference Card, Spanish	5806-AAR	1957794

Program Package Upgrade: A 3.5-inch program diskette and CD-ROM with documentation.

Description	Order Type	Part Number
VoiceType Dictation for Windows 1.32		
Software Upgrade, U.S. English	5803-AAR	52H3640
Software Upgrade, Spanish	5803-AAR	1957862

Program Upgrade

The program package upgrade option allows currently licensed users of VoiceType Dictation for Windows 1.1 and 1.1.1, or VoiceType Dictation for OS/2 1.1, or IBM Personal Dictation System (IPDS™) 1.0, to obtain the new function for a program upgrade charge. The program upgrade to VoiceType Dictation for Windows 1.32, U.S. English is available when the base product becomes available. Orders will be accepted now. These upgrades will be available through all of the same channels that the product is being sold through. The remarketer will require a Proof of License for each upgrade ordered. The original Proof of License is the separate Proof of License page in the program package as identified by IBM to the remarketer.

Customers can also order upgrades through the local IBM office using the IBM Personal Computer/System Program Upgrade and Certification Order Form.

Terms and Conditions

Licensing: The program described in this announcement is licensed under the terms and conditions of the IBM International Program License Agreement. A copy of the agreement is included in the program package.

Variable Charges Apply: No

Limited Warranty: Programs acquired by the customer during the program services period will have a one-year limited warranty.

Program Services: Program services will be available until December 31, 1996. The statement of service included with the program describes this service.

Presale information, installation, and technical support is provided by TALK-2-ME. TALK-2-ME can be reached via CompuServe Information Service (GO TALK2ME), the Internet (Marketing = talk2me@bcrvm1.vnet.ibm.com or Technical Support = talk2me@dalvm41b.vnet.ibm.com), or by calling 800-TALK-2-ME (825-5263). Support is available Monday through Friday from:

- Presale: 8:00 a.m. to 6:00 p.m. EST
- Postsale: 7:00 a.m. to 7:00 p.m. CST

Volume Orders: Not applicable

Authorization for Copy and Use on Home/Portable Computer: Yes, as specified in the program package

Program Transfer: Fully transferable

Educational Allowance: A reduced charge is available to qualified education customers. The educational allowance may not be added to any other discount or allowance.

Charges

Description	Order Type	Feature Number	Part Number	OTC*
VoiceType Dictation for Windows 1.32, U.S. English				
Software only	5801-AAR	0749	52H3636	$ 599
Software only Upgrade	5803-AAR	0269	52H3640	99
ISA Solution Pac	5801-AAR	0750	52H3630	999
MCA Solution Pac	5801-AAR	0751	52H3632	1,099
PCMCIA Solution Pac	5801-AAR	0752	52H3634	1,099
VoiceType Dictation for Windows 1.32, Spanish				
Software only	5801-AAR	1138	1957791	599
Software only Upgrade	5803-AAR	0283	1957862	99
ISA Solution Pac	5801-AAR	1140	1957855	999
MCA Solution Pac	5801-AAR	1141	1957858	1,099
PCMCIA Solution Pac	5801-AAR	1142	1957860	1,099

Program Upgrade Charge: A program upgrade charge will apply for upgrades from VoiceType Dictation for Windows 1.1 and 1.1.1 to VoiceType Dictation for Windows 1.32.

*   One-Time Charge

**   Company, product, or service name may be a trademark or service mark of others.

®   Registered trademark of International Business Machines Corporation