# Table of Contents

- Disclaimer .................................................................................................................. 3
- About the Vocal Joystick ............................................................................................. 4
  - Introduction ............................................................................................................... 4
- Overview of How to Get Started .................................................................................. 4
- Quick Tutorial .............................................................................................................. 5
  - Making the pointer move ......................................................................................... 5
  - Changing directions and speed ................................................................................. 5
  - Do not yell .................................................................................................................. 5
  - How to click ............................................................................................................... 6
- For Further Information ............................................................................................... 6
- Prerequisites ................................................................................................................ 7
- Windows XP/Vista ....................................................................................................... 8
  - Application Overview ............................................................................................... 8
  - Installation ................................................................................................................ 10
  - Creating a User Profile ............................................................................................. 13
  - Using the Vocal Joystick ......................................................................................... 16
  - Troubleshooting ...................................................................................................... 18
- Linux .......................................................................................................................... 20
  - Application Overview ............................................................................................... 20
  - Installation ................................................................................................................ 22
  - Creating a User Profile ............................................................................................. 22
  - Using the Vocal Joystick ......................................................................................... 25
  - Troubleshooting ...................................................................................................... 26
Disclaimer

Legal warnings and copyrights

This is a beta release of experimental research prototype. It is being released for entertainment purposes only and should be neither used nor relied upon for any medical or rehabilitative purpose.

The program is made available “AS IS” with no warranty expressed or implied.

We cannot guarantee any technical support for the program, and may not be able to respond to any email requests for support. We do welcome any comments/feedback you may have.

As an additional resource for connecting with other users who are using the Vocal Joystick, we have set up an online forum where you can post questions or comments about your experience at http://vocaljoystick.createmybb2.com/

We have done our best to ensure the correctness of our software, but we assume no responsibility for damage caused by our program to you or your computer.

We assume no responsibility or liability for, and provide no guarantees against, any physical or mental effects that may be in any way attributable to excessive use of your voice or the Vocal Joystick system, including but not limited to soreness, redness, damage, or inflammation of any kind to your voice, throat, mouth or any other part of your body or your surrounding environment.

The Linux version of this software uses the FOX Toolkit Library (http://www.fox-toolkit.org), which is covered by the Lesser GNU Public License and a FOX-specific addendum.
About the Vocal Joystick

Introduction

The Vocal Joystick is a software tool developed at the University of Washington, Seattle that enables you to control the mouse pointer on your computer using your voice.

Unlike conventional approaches to speech-driven pointer control, such as issuing directional commands (e.g., “Move mouse upper right... faster... stop. Click.”) or grid-based commands (e.g., “4...7...2...Click”), the Vocal Joystick provides a way to smoothly control the mouse pointer on a computer screen by using continuous vocalization of vowel sounds (e.g., “aaeeeeoooo”) and a small set of consonant sounds (e.g., “ch”, or “k”) to cause button clicks.

Overview of How to Get Started

In order to start using the Vocal Joystick for mouse control, you need to follow the following 4 steps:

1) **Download and install** the Vocal Joystick software from our web page: http://www.vocaljoystick.org
2) **Train the Vocal Joystick software** to learn aspects of your voice (this only takes a few minutes).
3) **Run the Vocal Joystick mouse controller**.
4) Start having **fun!! 😊**

Each of the steps may be slightly different depending on the operating system of your computer, so please follow the detailed instruction in the following sections for your relevant operating system.

The following Quick Tutorial section explains how to actually control the mouse pointer once you get the Vocal Joystick installed on your computer.
Quick Tutorial

Making the pointer move
To make the pointer move in a particular direction, all you need to do is just utter the vowel sound corresponding to that direction, as shown in the red font in the figure below. You only utter the red vowel sounds and not the whole word, so to move the pointer up, instead of uttering “cat”, you would just utter the “a” sound in the word “cat”. Just like a real joystick, the pointer will continue moving while you vocalize, and will stop moving as soon as you stop vocalizing.

Changing directions and speed
You can change the vowel sound during vocalization to change directions. For example, if you start out with the “a” sound (from the word “cat”) and then slowly change the sound to the “ee” sound (from the word “feet”), the pointer will start out moving up and then will start moving to the left. The faster you change from “a” to “ee”, the faster the direction of movement will change from up to left. Also, you can control how fast the pointer moves – that is, the pointer will move faster the louder you vocalize, and slower the softer you vocalize.

Do not yell
In general, you should not need to speak loudly when you use the Vocal Joystick – we have designed it to work using a close-talking microphone but where you only
need to utter the above sounds pretty quietly. This is also good if you plan to use the Vocal Joystick for many hours.

**How to click**
To simulate a left mouse button click, you make a short “ck” sound. You can also toggle the mouse button state (e.g., for dragging) by making the “ch” sound.

**For Further Information**
To learn more about the technology behind the Vocal Joystick, view demonstration videos, read published academic papers, and see a list of related projects, please visit the project website at http://www.vocaljoystick.org.
Prerequisites

In order to install the Vocal Joystick, you need the following setup:

**Software:**
- Operating System: any of the following:
  - Windows XP SP3 (32-bit)
  - Windows Vista SP1 (32-bit)
  - Linux (32-bit, built on Ubuntu 8.04 LTS – Hardy Heron)
- *For Windows XP/Vista users:*
  - .NET Framework 2.0 or higher
    (you can download the current version of the .NET Framework from http://www.microsoft.com/net/DownloadCurrent.aspx)
- *For Linux users:*
  - X11 and X11tst libraries
  - sox library with support for .wav files

**Hardware:**
- CPU: 2.0GHz or faster
- Memory: 512MB or more
- Microphone: USB (strongly suggested) or mini plug close-talking microphone. While the Vocal Joystick might work with a laptop microphone, we have only tested the current system with close-talking microphones.
Windows XP/Vista

Application Overview

The first thing you must do before starting to use the Vocal Joystick is to train the system to learn aspects of your voice. Technically speaking, this process is called “adaptation,” since you are adapting the Vocal Joystick’s internal algorithms so that they will work better with your voice.

The following figure shows the main parts of the Vocal Joystick user interface and their functions:

1. **Number of Vowels** – The Vocal Joystick can respond to either 4 or 8 primary directions. Using 4 vowels is better when you first start using the Vocal Joystick since it gives you very precise control of a small set of directions, but it makes it more difficult to move in diagonal directions. Once you have become proficient at the 4-vowel version of the Vocal
Joystick, you can try using 8 vowels to allow greater flexibility in directional control.

2 **User Profile** – Click *New* to create a new user profile or select one from the dropdown list. User profiles may be used to store what the Vocal Joystick has learned about your voice, so that you do not need to run the training module again in the future. However, any time your voice or surrounding environment has changed (say you have a cold, you have laptop and have moved locations, there is noise in the background like a bathroom fan, or you have changed to a different microphone), it is better to run the training module again.

3 **Sound Samples** – Click on a speaker icon next to each vowel and discrete sound buttons to hear how they should be pronounced. The figure above shows 8 speaker icons for the vowels. If you click on the 4 vowel radio button, only 4 speaker icons will be shown.

4 **Vowel Display Mode** – When checked, displays the vowel sounds as International Phonetic Alphabet (IPA) symbols instead of English words. To learn more about IPA, see the following link: http://en.wikipedia.org/wiki/International_Phonetic_Alphabet

5 **Vowel Recording Buttons** – Click each button to record your voice for each vowel for training. A blue color indicates vowels that you have already recorded, and an orange color indicates ones which are not yet recorded.

6 **Discrete Sound Recording Buttons** – Similar to the Vowel Recording Buttons except for discrete sounds (non-vowel sounds) which are used for mouse clicks.

7 **Train Button** – After all the necessary vowels and discrete sounds have been recorded (all buttons are blue), click this button to tell the system that it should use the samples you have given to learn about your voice.

8 **Run Vocal Joystick Button** – After the *Train Button* has been clicked, click this button to start using Vocal Joystick to control the mouse pointer.
Installation

Follow the steps below to install Vocal Joystick on your computer.


2. Double click on VocalJoystick.msi. In the dialog that appears, click Next.

![Welcome to the Vocal Joystick Setup Wizard]

The installer will guide you through the steps required to install Vocal Joystick on your computer.

WARNING: This computer program is protected by copyright law and international treaties. Unauthorized duplication or distribution of this program, or any portion of it, may result in severe civil or criminal penalties, and will be prosecuted to the maximum extent possible under the law.
3. Choose the directory where you would like the program to be installed, and whether to make the Vocal Joystick program available for all users or just the current user, and click Next.
4. Click Next to start the installation.  

**Note to users running Windows Vista:** If a User Account Control dialog appears warning that an unidentified program wants access to your computer, click *Allow*.

5. Once the installation is complete, click *Close*.  

Please use Windows Update to check for any critical updates to the .NET Framework.
Creating a User Profile

Before using the Vocal Joystick, a new user profile needs to be created for the system to be able to recognize your voice.

1. Place a close-talking microphone on your head and then plug the microphone into the USB port of the computer. This next step is important to verify that the microphone is recording properly: Make sure that your recorded speech from the microphone is not distorted in any way. When you speak normally, the recorded speech should sound clear, crisp, and distortion-free. If there is distortion, adjust the sound properties until the speech is clear. We are not in the position to diagnose or help with any sound hardware or software problems you might be having. There are many resources on the Web that can help you set up a closed-talking microphone with your computer.

2. Open the Vocal Joystick Control Panel by opening the Start menu and selecting All Programs -> Vocal Joystick -> Vocal Joystick.
3. Click on the *New* button next to the *Username* dropdown box, and type in your name.
4. **IMPORTANT**: following this next step very precisely is **crucial** in order for the Vocal Joystick system to work. For each of the vowel sounds shown in the vowel diagram on page 8, take a breath and then click on the corresponding button to record a sample of your voice. A black window will pop up similar to the one shown above. Remain quiet when you click on the button, and once you see a line that reads “Noise power:” verify that it is less than 100. If it is higher, wait a few seconds while remaining quiet, and another line should appear reading “Noise power:” If that number is still larger than 100, close the window by clicking on the “X” button on the upper right corner, and click on the vowel button again. Once again, the summary of the process for this step are as follows (visit the Vocal Joystick download webpage to see a video of this process):

1. Stay quiet  
2. Click button  
3. Wait for "Noise power"  
4. Say the sound
5. If you have clicked on a vowel button, continue uttering the vowel sound for about 2 seconds at a normal loudness, until the black window goes away. Try not to stop vocalizing before the black window goes away, but if you run out of air, take a breath and resume uttering the vowel sound. Again, as mentioned in the previous step, it is advisable to take a breath before clicking on the vowel sound buttons.

If you have clicked on a discrete sound button, you need to give only one short example of such sound (less than half a second). Repeat the same procedure for the two discrete sounds “k” and “ch.” Try not to have any vowel sound following the discrete sound, such as “chuh”.

6. Once all sounds are recorded, click on the *Train* button. A black window appears with various texts scrolling by. Wait until it disappears.

**Using the Vocal Joystick**

1. Once you have created a user profile and clicked on the *Train* button, click on the *Run Vocal Joystick* button to start controlling the mouse cursor using your voice. A black window appears.
2. You can now control the mouse pointer by uttering the vowel sounds shown in the vowel map figure in the Quick Tutorial section.

3. If at this point you want to minimize the black window so that you can see the rest of the screen, you may do so.

4. To stop using Vocal Joystick to control the mouse pointer, close the black window that appeared after step 1 above by clicking on the red X button on the upper right corner of the window. This may even be accomplished using the Vocal Joystick to control your cursor!
Troubleshooting

When I click on a vowel button or a discrete sound button to record, the “Initial noise power” number is always high and does not come down.

Try opening the Sound control panel, and decreasing the volume of your microphone. We are not in the position to help diagnose any problems with sound hardware or software. Please see the following web page for help with these issues:


When I click on a vowel or a discrete sound button to record the sound, the black window does not go away after the text stops scrolling.

Occasionally, the black window main remain open even after you have finished uttering the vowel sound or the discrete sound, displaying “Stopping Frontend...” on the last line. If this happens, manually close the black window by clicking on the red “X” at the top right corner of the window, and rerecord the sound.

When I say the vowel sounds, the mouse pointer moves too slow/fast.

The loudness at which you recorded the vowel sounds during the training stage determines the loudness at which the pointer will move at a normal speed. If you feel the mouse pointer is moving too fast, try re-running the training module again and re-recording the vowel sounds a little louder than before. If you feel the mouse pointer is moving too slowly, try re-recording the vowel sounds a little softer than before.

When I say the vowel sounds, the mouse pointer does not move in the right direction.

Try re-running the training module again. Make sure that the sound environment (background noises, type of microphone used, if you have a cold or not, room
within which your laptop resides) is exactly the same when you use the Vocal Joystick as when you adapted your voice to it.
Application Overview

The first thing you must do before starting to use the Vocal Joystick is to train the system to learn aspects of your voice. Technically speaking, this process is called “adaptation,” since you are adapting the Vocal Joystick’s internal algorithms so that they will work better with your voice.

The following figure shows the main parts of the Vocal Joystick user interface and their functions:

1. **Number of Vowels** – The Vocal Joystick can respond to either 4 or 8 primary directions. Using 4 vowels is better when you first start using the Vocal Joystick since it gives you very precise control of a small set of directions, but it makes it more difficult to move in diagonal directions. Once you have become proficient at the 4-vowel version of the Vocal

2. **User Profile**

3. **Vowel Recording Buttons**

4. **Discrete Sound Recording Buttons**

6. **Run Vocal Joystick Button**
Joystick, you can try using 8 vowels to allow greater flexibility in directional control.

2 **User Profile** – Enter a name in the black and click *Add New User* to create a new user profile or select one from the dropdown list. User profiles may be used to store what the Vocal Joystick has learned about your voice, so that you do not need to run the training module again in the future. However, any time your voice or surrounding environment has changed (say you have a cold, you have laptop and have moved locations, there is noise in the background like a bathroom fan, or you have changed to a different microphone), it is better to run the training module again.

3 **Vowel Recording Buttons** – Click each button to record your voice for each vowel for training. A blue color indicates vowels that you have already recorded, and an orange color indicates ones which are not yet recorded.

4 **Discrete Sound Recording Buttons** – Similar to the Vowel Recording Buttons except for discrete sounds (non-vowel sounds) which are used for mouse clicks.

5 **Train Button** – After all the necessary vowels and discrete sounds have been recorded (all buttons are blue), click this button to tell the system that it should use the samples you have given to learn about your voice.

6 **Run Vocal Joystick Button** – After the *Train Button* has been clicked, click this button to start using Vocal Joystick to control the mouse pointer.
Installation

Follow the steps below to install Vocal Joystick on your computer.


2. Unzip the file into your home directory, or some other location (henceforth parent_dir).

Creating a User Profile

Before using the Vocal Joystick, a new user profile needs to be created for the system to be able to recognize your voice.

1. Place a close-talking microphone on your head and then plug the microphone into the USB port of the computer. This next step is important to verify that the microphone is recording properly: Make sure that your recorded speech from the microphone is not distorted in any way. When you speak normally, the recorded speech should sound clear, crisp, and distortion-free. If there is distortion, adjust the sound properties until the speech is clear. We are not in the position to diagnose or help with any sound hardware or software problems you might be having. There are many resources on the Web that can help you set up a closed-talking microphone with your computer.

2. Open a terminal window and change directories to parent_dir/vj/vjapp/bin

3. Run ./controlpanelfox.sh to launch the Vocal Joystick Control Panel.
4. IMPORTANT: following this next step very precisely is **crucial** in order for the Vocal Joystick system to work. For each of the vowel sounds shown in the vowel diagram on page 8, take a breath and then click on the corresponding button to record a sample of your voice. You will see some text appear in the console window, similar to what you see above. Remain quiet when you click on the button, and once you see a line that reads “Noise power:” verify that it is less than 100. If it is higher, wait a few seconds while remaining quiet, and another line should appear reading “Noise power:” If that number is still larger than 100, close the window by clicking on the “X” button on the upper right corner, and click on the vowel button again. Once again, the summary of the process for this step are as follows (visit the Vocal Joystick download webpage to see a video of this process):

1. Stay quiet  
2. Click button  
3. Wait for "Noise power"  
4. Say the sound
5. If you have clicked on a vowel button, continue uttering the vowel sound for about 2 seconds at a normal loudness, until you see text like the following in the console window:

```
Writing out data...
Data written
Stopping Frontend
Frontend stopped
====================================
Recording complete
```

Try not to stop vocalizing until you see “Recording complete,” but if you run out of air, take a breath and resume uttering the vowel sound. Again, as mentioned in the previous step, it is advisable to take a breath before clicking on the vowel sound buttons.

If you have clicked on a discrete sound button, you need to give only one short example of such sound (less than half a second). Repeat the same procedure for the two discrete sounds “k” and “ch.” Try not to have any vowel sound following the discrete sound, such as “chuh”.

6. Once all sounds are recorded, click on the *Train* button. More text will scroll by in the console window. Wait until you see the line “Adaptation complete.”
Using the Vocal Joystick

1. Once you have created a user profile and clicked on the Train button, click on the Start VJ button to start controlling the mouse cursor using your voice. More text will appear in the console window. Again, check the noise power level and wait if it is too high.

2. You can now control the mouse pointer by uttering the vowel sounds shown in the vowel map figure in the Quick Tutorial section.

3. If at this point you want to minimize the console window so that you can see the rest of the screen, you may do so.

4. To stop using Vocal Joystick to control the mouse pointer, the Start VJ button has renamed itself Stop VJ. Simply click that button to stop the system.
Troubleshooting

When I launch the Control Panel, I get an error that my system cannot open the audio device (/dev/dsp1).

By default, the Vocal Joystick looks for a microphone on /dev/dsp1. If you need to specify another audio input device, you can provide an option when launching the control panel. For instance, if your microphone is on /dev/dsp2, simply run 

```
./controlpanelfox.sh -d /dev/dsp2.
```

When I click on a vowel button or a discrete sound button to record, the “Initial noise power” number is always high and does not come down.

Try opening the Sound control panel, and decreasing the volume of your microphone. We are not in the position to help diagnose any problems with sound hardware or software. Please see the following web page for help with these issues:

When I say the vowel sounds, the mouse pointer moves too slow/fast.

The loudness at which you recorded the vowel sounds during the training stage determines the loudness at which the pointer will move at a normal speed. If you feel the mouse pointer is moving too fast, try re-running the training module again and re-recording the vowel sounds a little louder than before. If you feel the mouse pointer is moving too slowly, try re-recording the vowel sounds a little softer than before.

When I say the vowel sounds, the mouse pointer does not move in the right direction.

Try re-running the training module again. Make sure that the sound environment (background noises, type of microphone used, if you have a cold or not, room
License Agreement

Basic idea:

1. The Vocal Joystick is not a medical device and should not be used as one.

2. Excessive use of your voice may cause injury. Take care to limit use.

3. This is a beta release of research software, not a finished product.

4. The technology is not perfect, and unintended mouse clicks may happen. Be careful that this does not have unintended consequences.

5. We provide no warranty and hold no liability associated with your use of the Vocal Joystick.

6. We retain all rights to the software, Vocal Joystick name and logo.

Our license:

Please read the following terms and conditions carefully before using the Vocal Joystick ("the software"). Your use, distribution or installation of this copy of the Vocal Joystick indicates your acceptance of this License. If you do not agree to any of the terms of this License, then do not install or use the Vocal Joystick.

This beta release of the Vocal Joystick system is an experimental research prototype being distributed for entertainment purposes. It should neither be used nor relied upon for any medical or rehabilitative purpose.

The Vocal Joystick, all accompanying files, data and materials, are distributed "AS IS" and with no warranties of any kind, whether express or implied, to the extent permitted by applicable law. The user must assume all risk of using the program. This disclaimer of warranty constitutes an essential part of the agreement.

To the extent permitted by applicable law, the members of the Vocal Joystick team and the University of Washington assume no liability for, and provide no
guarantees against, any damages of any sort associated with any use of the Vocal Joystick. This includes, but is not limited to, physical effects or mental effects that may be in any way attributable to excessive use of your voice or the Vocal Joystick system such as soreness, redness, damage or inflammation of any kind to your voice/vocal tract or any other part of your body.

The Vocal Joystick may not be redistributed; users must download their own copies directly from http://www.vocaljoystick.org. If you have multiple systems on which you want to use the Vocal Joystick, you may freely copy your own version between them.

In addition, in no event are you authorized to use the Vocal Joystick in applications or systems where the Vocal Joystick's failure to perform can reasonably be expected to result in a financial loss, physical injury, or in loss of life. Any such use by you is entirely at your own risk, and you agree to hold the Vocal Joystick team and the University of Washington harmless from any claims or losses relating to such unauthorized use.

The Vocal Joystick team is the owner of the copyright of the Vocal Joystick. All of its derivatives, title and accompanying materials are the exclusive property of the Vocal Joystick team. All rights of any kind, which are not expressly granted in this License, are entirely and exclusively reserved to and by the Vocal Joystick team. You may not rent, lease, modify, reverse engineer, de-compile, or disassemble the Vocal Joystick.

You agree to indemnify and hold the Vocal Joystick team and the University of Washington harmless from any claims resulting from the use of the Vocal Joystick which may damage any other party.