UTAU: The Free Digital Singer
An Introductory Guide to Creating Synthetic Vocal Tracks in UTAU

Written by Cdra
Revision 1.0 (released 12/10/2012)

The voice is a powerful musical instrument. Lyrics add a new dimension to instrumental tracks, giving it more power and emotion. But the human voice, despite being a powerful instrument, is one of the most difficult to use. Getting the perfect vocals for a track can be at worst impossible, especially in today's world of electronic music. People who can't sing use autotune to correct themselves, and others still use it to make themselves blend with their electronic-styled music better. But what if you could have complete control of the voice? You could select every sound, every note, place and move them as you see fit—like any other synthesized instrument? That technology is available in the form of a free singing synthesizer application called UTAU.

Contents:

Background

In order to understand UTAU, it is necessary to understand its predecessor, VOCALOID. VOCALOID is a professional, sample-based vocal synthesis program developed by Yamaha; it debuted in 2004. The first VOCALOID engine, referred to as "Vocaloid1" by fans, was released first with two English voices. Later, the program was released with Japanese voicebanks; the Japanese culture took very well to VOCALOID, particularly after the release of "Hatsune Miku" on the "Vocaloid2" engine. An extensive fanbase has since developed for the software, its avatars, and the music produced with it. The engine is now on its third iteration and there are more than fifty available voices in various languages.

UTAU was written by Ameya, a Japanese man who wished to create a free alternative to the rather expensive VOCALOID software. It was first released in March 2008, making its claim to fame with a prank "fake VOCALOID" released on April Fool's Day of the same year. By appealing to the established VOCALOID fanbase, the program immediately settled into a niche of Japanese users. In time, the interface was translated into English, opening the way for English-speaking users to use this program. Though it is designed for use with the Japanese language, many users have since found methods to make UTAU emulate other languages with increasingly promising results.

The goal of this guide is to show you how to use PC UTAU, regardless of your interest in it— whether you want to produce original songs, cover songs, or just muck around in it. Some basic knowledge of music theory, such as pitch notation and piano roll notation, will be beneficial to learning to use UTAU. If you don't have that, don't worry; it's simple enough to pick up without any prior knowledge at all.

If you are already familiar with the VOCALOID program, you will find that UTAU operates in a very different way, and very little of your knowledge (beyond tuning theory) will transfer between the programs. Thus, this guide will be helpful to VOCALOID users as well.

Preparing to Install UTAU

In order to run UTAU, you will need a computer with the following specifications:

	Minimum	Reccommended
Operating System	Windows 2000	Windows XP Professional SP2 and newer
CPU	Intel Celeron Class CPU, 2 GHz	At least 2 GHz; Intel Core CPU, AMD (Sempron, Athlon, Opteron, Phenom) or better. Multicore preferred.
RAM	512 MB	1 GB RAM or more. 2 GB results in faster rendering in addition to better CPU.
Soundcard	Any	Any high-definition soundcard.
Disk Space	100 MB	The more space the better, as voicebanks take up plenty.

Table Data from UTAU wiki[1]

---

[1] http://utau.wikia.com/wiki/UTAU_wiki:System_Preparation

(As a note, there is a Mac version of UTAU called UTAU-Synth; however, as I only have Windows, I can't help you with how to install and use it. You will need to refer elsewhere for that information.)

It is also recommended, but not required, that you change your computer's System Locale to Japanese before installing. The main reason for this is that you will not be able to import MIDI files into UTAU if you do not, as the default voicebank's (it comes with the program) filenames will become nonsense ("mojibake") instead of Japanese characters. As your program will immediately assign the default voicebank to any midi you import, it will result in an error. You can change the default voicebank to one without Japanese filenames to avoid this issue.

You will also not be able to use voicebanks recorded with Japanese characters in their filenames if you do not change your System Locale to Japanese; thus, I recommend changing your locale if you intend to use Japanese-language voicebanks.

If you wish to change your system locale, you should refer to the guides on Microsoft's support website[2]. If you are using a shared computer or otherwise cannot change your system locale, you may be able to use AppLocale to emulate the Japanese system locale. I do not have information on how to do this, so you will again need to refer elsewhere for how to do that. If you cannot change your locale at all, don't worry; many users are able to use UTAU just fine without ever changing their system locale or using kana-encoded voicebanks.

How to Install UTAU

1. Download the latest version of the UTAU program from Ameya's site, http://utau2008.xrea.jp/. At the time of writing, the latest version of the program is UTAU v0.4.12; simply click the link that says "v0.4.12 インストーラ" to download the installer.
2. Open the downloaded zipped file and run the installer inside, "utau0412installer.exe."
3. The installer is in Japanese, but don't worry—it's not much different from any other program you've installed. The button at the bottom right-hand corner that says "N" will take you to the next screen, the one with "B" to its left will take you back a screen, and the one to the left of "B" will cancel the installation. Click "N" to proceed.
4. You may select the directory you wish to install the program to. Automatically, it will install to your 32-bit (x86) Programs folder. Click the button to the right that says "R" to search for a new directory if you wish. Below the directory selection bar are two options, "E" and "M." "E" will install the program for all users on your machine; "M" will install it only for your current user profile. Once you've changed what you want to change, click "N" again to proceed.
5. Click "N" again to begin the installation process. Give it some to complete; it could take a few minutes.
6. When it finishes, a final dialog will appear. Click "C" to close the installer.
7. Before you run UTAU, go ahead and patch the program into English. The patch can be found at Mianaito's site, http://www.voiceblog.jp/mianaito/1062049.html. Simply follow the instructions on this page to apply the English patch.
8. Congratulations, your UTAU program should now be installed and run in English! If you encounter any errors in running it, you should check the troubleshooting guide on the UTAU wiki, on the "Downloading and Installing UTAU" page[3].

---

2 http://windows.microsoft.com/en-US/
3 http://utau.wikia.com/wiki/UTAU_wiki:Downloading_and_Installing_UTAU

The Synthesis Window: An Overview

You'll find the piano roll layout of UTAU familiar if you have used a synthesizing program in the past. The representation of the pitch appears on the piano roll (along the far left) when you hover over the corresponding line anywhere else in the window; as placeholders, the C pitches are shown at all times.

4

Render

ACPT P2P3 P1P4 OPT RESET

3   4   5

Icon Palette

Selected Region (gray)

Scroll bar (vertical, through piano roll) →

otes (pink)

55
ld R

Note

–R

–R

–R

99
wa

92
–u–

R

Unselected notes (blue)

Zoom Vertical
Zoom Horizontal

ar (horizontal; through song)   Move Multiple

M
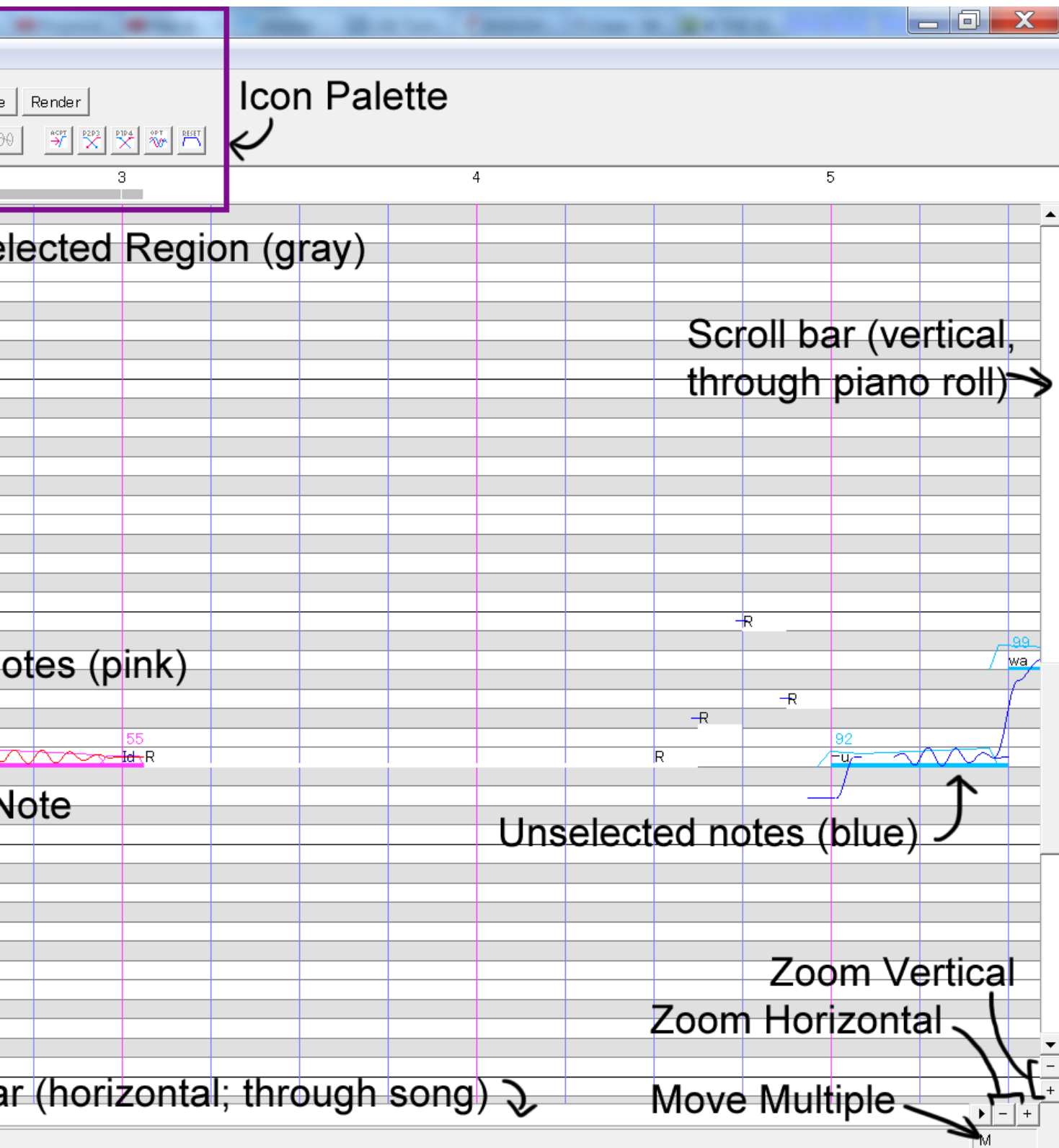
In UTAU, C4 represents "middle C." You can see the piano keys corresponding to the pitches on the far left.

The Icon Palette

The icon palette contains many of the program's most important functions, making it very complex. The buttons are labeled and explained below:
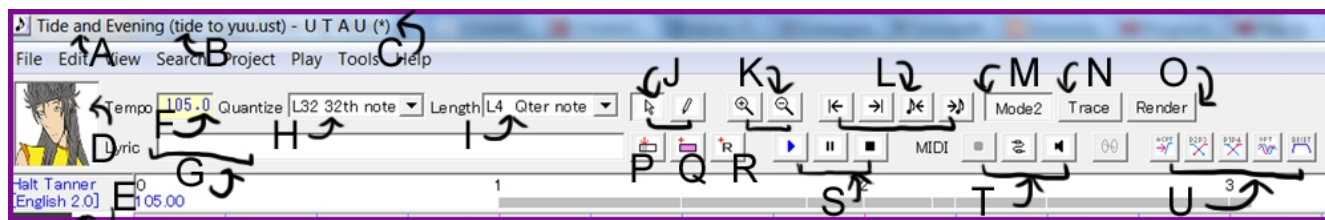

Illustration 1: The icon palette.

A) Project Title- the name of the song you're working on, usually.

B) File Name- The name of the file you are working on; .ust is the extension for UTAU files.

C) This asterisk appears if you have changed the file since you last saved it.

D) Voicebank Icon- The avatar of the voiceabnk is represented here.

E) Voicebank Name- Click this box to open the Project Properties dialog (see "Voicebanks").

F) Tempo- Click this box to enter your desired project tempo.

G) Lyric Bar- Enter lyrics into this bar to use with icons P and Q.

H) Quantize- Determines the minimum amount by which you can change the length of a note. Your note lengths are rounded to the nearest "quantize" value when you change them.

I) Length- The default length of a note or rest when it is created using icon Q or R respectively.

J) Tool Select- You may either use the arrow tool (drag and select notes) or the pencil tool (draw notes onto the piano roll).

K) Zoom Tools- Zoom in with the + magnifying glass, out with the – magnifying glass. The zoom affects the horizontal scale of the window.

L) Quick Move Buttons- The first two buttons take you to the beginning of the file and the end of the file respectively. The latter two take you to the first and last non-rest note in the file respectively.

M) Pitch Editor Mode Select- When this button is pressed, you are in Mode2 pitch editing. You should stay in Mode2 pitch editing at all times. You may refer to Appendix A: "Mode1 vs Mode2 Pitch Editing" to learn more about why this is.

N) Trace- Makes any Mode1 pitch edits visible in Mode2 as gray lines, so that they can be traced.

O) Render- Renders any Mode2 pitch edits into Mode1. I don't find this particularly useful.

P) Replace Lyrics- Replaces lyrics in the selection with those entered in the lyrics bar, (G).

Q) Insert Lyrics- Inserts new notes of the length specified in (I) just prior to the selection, with the lyrics entered in the lyrics bar (G).

R) Insert Rest- Inserts a rest of the length specified in (I) just prior to the selection.

S) Play, Pause, and Stop Buttons- Self-explanatory; for playback.

T) MIDI Controls- Record MIDI input, connect to a MIDI controller, and monitor MIDI output in order. I will not discuss these functions in this tutorial, but perhaps in a later one.

U) Automatic Envelope Controls- The buttons found here are ACPT, P2P3, P1P4, OPT, and RESET. They will be discussed in greater detail in the Tuning section of the guide.

    • ACPT- Locks the note's timing parameters according to the voicebank configurations.

    • P2P3/P1P4- Automatically crossfade the notes at the 2nd and 3rd points of the envelope or 1st and 4th points of the envelope respectively. This crossfade is necessary for smooth singing.

    • OPT- Short for "Optimize Crossfade." Uses the cache of the program to optimize the crossfades for a smoother result; implement it on a final render for smoother results.

    • RESET- Resets the envelope to its usual rectangular shape.

Notes

The lyric (entered as phonemes, phonetic representations of syllables) sits at the front of the note, and represents the sound that note will make. Japanese is a phonetic language, so the representations of the sounds are simple and intuitive; however, for more complex languages such as English, user-created phonetic systems are used to represent the sounds. In the example notes shown in Illustration 2, this phonetic system uses "-u" to represent the sound "uh" configured to follow a rest. The other phonemes, "dri" and "ift", are more self-explanatory. This series of phonemes sounds out the word "adrift."
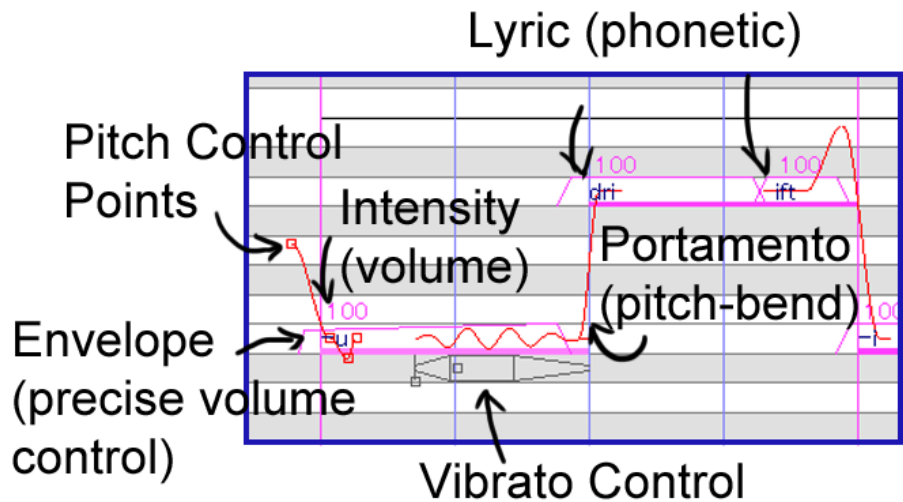

Illustration 2: Parts of the notes.

Notice how each syllable takes on its own note on the piano roll, and ending sounds (like "ift") are placed at the end of the note if needed. The vowel is the center of the note; there are consonants at the beginning and/or end of a note, but they are separated into a beginning and ending note as shown in illustration 2. Additionally, phonemes are case-sensitive. You can double-click the lyric to change it.

Japanese voicebanks that are either CV or VCV recorded will not have ending sounds like the "ift" in the example. Only CV VC voicebanks will have these sounds; you can find a more detailed discussion of how to use different voicebank styles in "Voicebanks."

Pitch control points control the Portamento (pitch connectivity)/"pitch-bend" functions. The vibrato control box comes from the same dialog window. This is the system used to edit pitch within a note; it will be discussed in-depth in "Tuning."

You can edit the intensity (volume) of a note by clicking on the value and dragging it up or down. The envelope, a finer volume control, serves two purposes: crossfading notes together for smoothness, and editing note dynamics (ie crescendo and decrescendo). You can edit it by opening the Envelope dialog, which will be discussed in-depth in "Tuning."

View Toggles
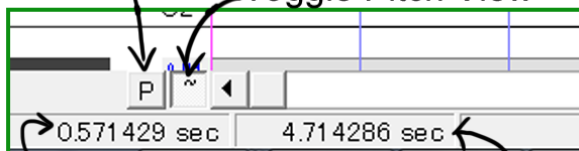
Illustration 3: The "view toggles."

In the bottom left-hand corner of the program window, you can see the length of a standard quarter note in your tempo (in seconds) and the length of the region you've selected (in seconds). This can be useful for seeing how long your rendered wav file will be.

You can also change the your view on the piano roll. If you toggle flag view ("show parameters"), the program displays note parameters below each note in red, as shown in illustration 4.

7

In the example, the only parameter in place is "mod0," which prevents the notes from going off-pitch. I will discuss parameters (flags) in-depth during "Tuning."
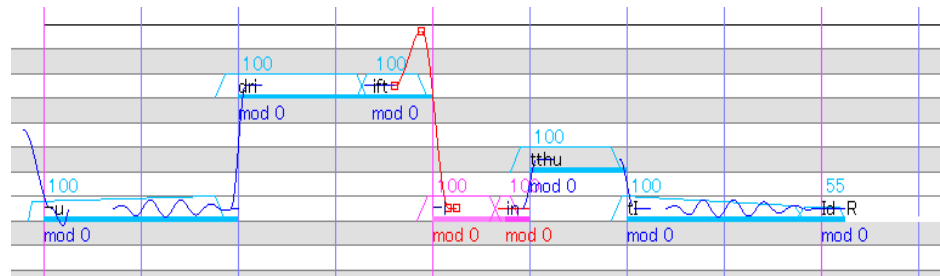

Illustration 4: Parameter view on.

If you click the [~], you will toggle off pitch view ("show pitch curves") and see the notes as plain blocks, as shown in illustration 5. The pitch edits are now hidden, though they are still audible. This is useful for MIDI entry, before you worry about tuning. However, it makes it hard to tell what you're doing, so be careful using it.
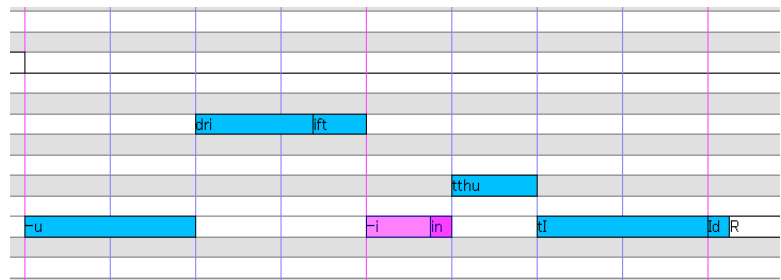

Illustration 5: Pitch view off.

Voicebanks

A "voicebank," also sometimes called a "voice library," is the collection of sounds used to create singing in UTAU. Voicebanks contain all of the phonemes required to replicate a given language; some may even be able to replicate more than one language. They are often large files, given that each voice sample is an uncompressed .wav audio file. When you download a compressed voicebank, you get a folder with the required audio files, ".frq" files for each wav file (contain the pitch data for that wav), the required configuration file(s) (oto.ini, for creating smooth, on-time singing; and, optionally, prefix.map, for placing the

Illustration 6: Project Properties Dialog

voicebank's recordings at the right pitches, if it was recorded at more than one pitch), and a readme file. Some may have other supplemental information, such as character reference art or additional, non-singing phonemes such as breaths.
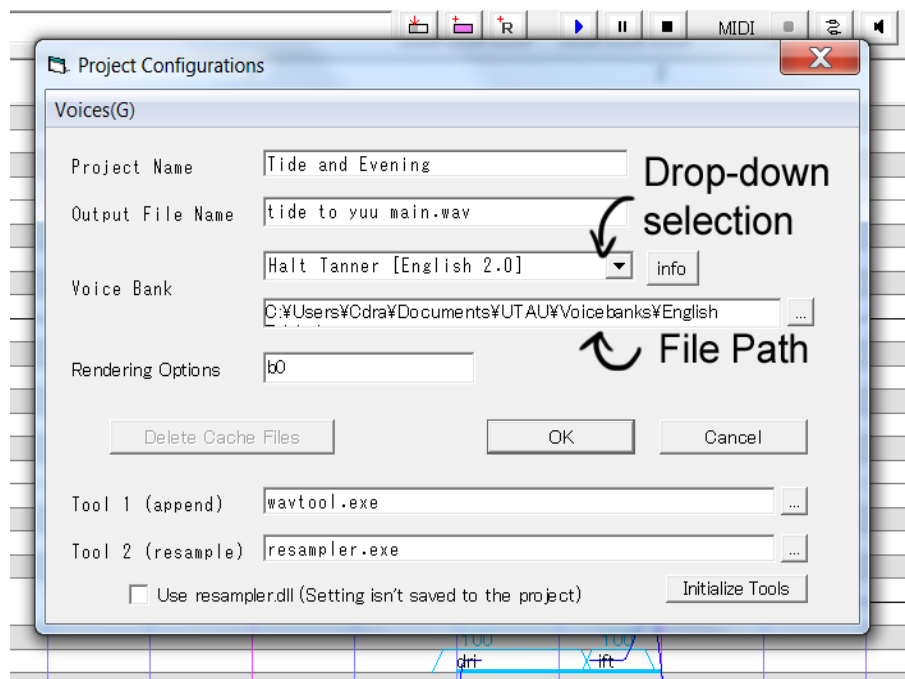
Most voicebanks are publicly distributed by individuals. They are free to download and use, but be sure to read the terms of use of a voicebank before using it.

There are thousands of voicebanks being distributed online; most of them are Japanese-language only due to the program's Japanese origin, but the number of non-Japanese voicebanks available

is growing. To begin a search for the perfect voice for your song, start with UTAU.me[4]. You can also check out the UTAU Wiki[5] and search for voicebanks there. Of course, you also may want to use your own voice in UTAU, as many people do; there are many tutorials online, particularly on YouTube, dedicated to helping you create a voicebank. I ask that you refer to those tutorials if you wish to create your own UTAU voice.

Once you find the voice you want, simply download it and extract it into the folder "voice," inside the folder where you installed UTAU. You will then be able to select the voicebank by name from the drop down window in the project properties dialog (illustration 6).

Since English is a very complex language with a very simple alphabet, people use many different phonetic systems to describe it. Every voicebank probably has its own variation of the system, so you should make sure to study it before you use the voicebank. Other languages, like Korean and Japanese, are phonetic, and therefore do not have this problem.

Voicebank Styles

Also of note when downloading (or creating) UTAU voicebanks is the voicebank style. There are three main recording styles used for voicebanks: CV, VCV, and CV VC. In each style name, "C" refers to a consonant and "V" refers to a vowel.

- CV: CV is generally considered the most basic style of voicebank, and is used for Japanese voicebanks. CV banks are generally small and easy to create, but they are limited; Japanese is one of the only languages that can be perfectly replicated using a CV voicebank, since it does not have any ending consonant sounds.
- VCV: VCV is a more advanced recording style that utilizes leading-in vowels on each note, such that a clean crossfade can be created between vowels. These voicebanks are much larger than CV banks, but generally have smoother singing results with more realistic note transitions. However, VCV is, like CV, not an effective method for languages other than Japanese, as it relies on vowel-based transitions that may not occur in the same way in many other languages. Additionally, Japanese has very few phonemes (required sounds) than most other languages; a VCV voicebank for another language is made less viable by the fact that almost all other languages have a lot more required phonemes than Japanese. Non-Japanese VCV banks would be extremely large.
- CV VC: CV VC is generally the best recording style for non-Japanese languages, as the "VC" represents its ability to handle ending consonants. The examples show previously in illustrations 2, 4, and 5 show a CV VC English voicebank. CV VC is also usable for Japanese; essentially it is recorded like VCV, then the VC sounds are used in between notes to create "manual VCV" sounds. CV VC editing will be covered in depth in Appendix B: "CV VC Editing."

Bear in mind the style of a voicebank when creating or editing a ust for it.

Sequencing Notes: How to Use the Piano Roll

You can enter notes to the piano roll either by typing the lyrics into the lyrics bar and using the insert lyrics button to add them to the piano roll, or by selecting the pencil tool and drawing the notes onto the piano roll. In either case, you will probably have to edit the lyric by double-clicking on it and typing in the phonetic representation. I will discuss importing MIDI and VSQ (vocaloid sequence) files to UTAU in Appendix C: "Importing MIDIs and VSQs."

---

4 http://utau.me/profile/
5 http://utau.wikia.com/wiki/UTAU_wiki

If you are using a premade .ust file, the notes will already be in place, but you may still need to edit their lengths.

When editing the length of a note, there are different key combinations you can use.

Keys	Shorten Note (drag left)	Lengthen Note (drag right)
Click	Shortens note and creates a rest of the length by which you shortened the note	Overwrites a rest in front of the note (or empty space); does not work if there is another note in front
Shift+Click	Shortens the note and shifts the notes that follow it back (left) by the same amount	Lengthens the note and shifts the notes that follow it forward (right) by the same amount
Ctrl+Click	Shortens the note and lengthens the next note by the same amount	Lengthens the note and shortens the next note by the same amount (can cause the next note to disappear)
Shift+Ctrl+Click	Splits the note in two; the leftmost note keeps all the properties, where the rightmost note loses them	Nothing

Use the correct combination for how you want to alter your notes! Shift+Ctrl+Click is especially useful for splitting notes for CV VC ust editing.

Entering the notes is a simple enough process, but the voice will sound choppy at first. Select the notes (you can use Ctrl+A to select all notes, like in most other programs) and click P2P3 (or P1P4 if you prefer; generally I think P2P3 creates less envelope errors). This will crossfade the lyrics so that they aren't so choppy.

If you're using a premade ust, you'll still need to crossfade the notes, but before you do so you should clear the parameters on all of the notes. Clearing parameters will be discussed in the next section, "Tuning."

However, even though the voice will no longer be choppy-sounding, it could sound very robotic and lack any feeling at all. But there are ways to make it more emotional using the various tuning tools at your disposal.

Tuning—The Voice as an Instrument

You wouldn't write a song with an untuned guitar, would you? Like any instrument, the mechanical voice will need to be tuned to sound good. Tuning can be used to imbue the voice with emotion and make it sound clearer and more human.

Envelopes

Envelopes are the detailed volume control within a note. Right-click on a single note (not a region) and click "Envelope" to open the envelope editing dialog. The pink line is your default intensity; placing control points above it will make that section louder, while placing control points below it will make the section quieter.

- Control Points (top bar): The location of the five (four shown) control points in the note. The first box for each ("p") shows the location in ms from the beginning of the note; the second box ("v") shows the volume (intensity).
- Visual Envelope Display: This is the easiest way to edit the envelope. Click and drag the control points to create crescendos, decrescendos, and other such dynamics in the note.

Illustration 7: The envelope editing dialog with unedited envelope.

- Pre/Ovl: Short for "Preutterance" and "Overlap," these are configuration values that determine the starting point of the note the note (with respect to the original voice sample) and the length of the crossfade (with respect to the preutterance) respectively. They have default values as given by the voicebank's configuration files, and should not be changed; if they have a value in them, you may have clicked ACPT on the note. These parameters will be discussed in extreme detail in a future tutorial.

Illustration 8: The envelope after P2P3 crossfade.

- Origin: Locks the default configured values of Preutterance and Overlap.
- Normal: Normalizes the envelope.
- Chain: If "Chain" is checked, the points of the envelope all move together; if you click one control point and move it, the others will compensate somewhat. If you uncheck it, they will move independently.
- Reset: Resets the envelope to the state shown in illustration 7.
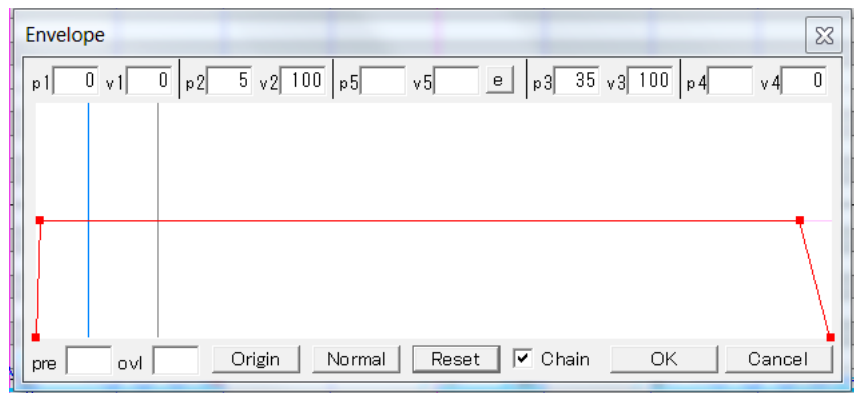

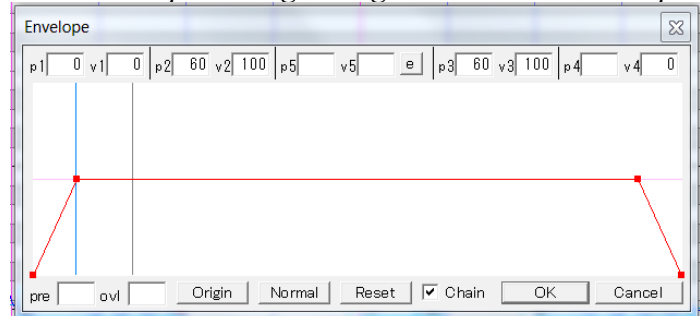*Illustration 9: The envelope after some editing.*

The envelope has five control points, though the fifth will only appear if you give a value in the "p5" box in the top center of the dialog. You will need to use either p1/p4 or p2/p3 as crossfade points to make the singing smoother, as mentioned before. The easiest way to do this is to select all of the notes, then click P1P4 (for p1/p4 crossfade) or P2P3 (for p2/p3 crossfade) in the Automatic Envelope Controls section of the Icon Palette.

Illustration 10: Envelope error—Notice how the control points are out of order.

Once the crossfades are complete, you can use the control points to create dynamics within the note. This envelope will give the note a slight crescendo, then the volume will taper off near the end. Note how the p5 point was added by giving it a value. Experiment with different envelope shapes to get the effect you want!
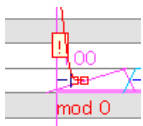
Illustration 11: Error.

If you see a red exclamation point in a box (as shown in illustration 10) above a note, that means there's an error with the envelope, usually where the control points are out of order (ie p3 is behind p2 due to the crossfade, etc). To fix this, open the envelope dialog and click "Normal" to normalize the envelope. If that still doesn't work, try dragging the control points out a little bit—take p2 to the left or p3 to the right, for instance—to clear up any crossed-over points.

Note Properties

The Note Properties dialog is the catch-all for miscellaneous vocal parameters in UTAU. To open this dialog, right click a selected note or region and click the last option, which will be "Property" or "Region Property" for a note or region respectively.

When selecting a region, gray boxes indicate that the value in that box varies over the region. You can also type a space into the gray boxes to clear them for the region. If you type anything into the boxes, it is set for the whole region.

Illustration 12: Note properties dialog

- Lyric: The phoneme contained in the note.
- Note and Length: The pitch and length of the note, respectively.
- Intensity: The base volume of the note.
- Modulation: Modulation, or "mod" for short, is best left at 0 to keep the singing on-pitch; high values of mod will cause the pitch to vary wildly, resulting in a "drunk" sound.
- Preutterance: Given a default value according to the configuration file for the voicebank. It is best not to do anything to this field unless it contains a value that you did not put there , in which case click the "Clear" button to remove it.
- Overlap: Like preutterance, overlap is given a default value according to the configuration file for the voicebank. Again, it is best not to do anything to this field unless it contains a value you did not put there using ACPT or otherwise; if this happens, click "Clear" to remove the value.
- Consonant Velocity: This determines the length of the consonant portion of the note. The vowel will be stretched to fit the note, but the consonant keeps the same length regardless. If your consonant seems to be cut off, increase the consonant velocity to make the consonant sound faster, which will allow it to fit into smaller note spaces. The default value is 100, the minimum value is 0, and the maximum value is 200.
- Show/Hide Details: The bottom section of the dialog can be collapsed by clicking "Hide." When it is collapsed, it will show "Details" instead, which can be used to show the section.
- BRE: BRE is short for "breathiness," and determines the amount of breath put into the voice. Using a high breathiness (greater than 50) will give the voice more
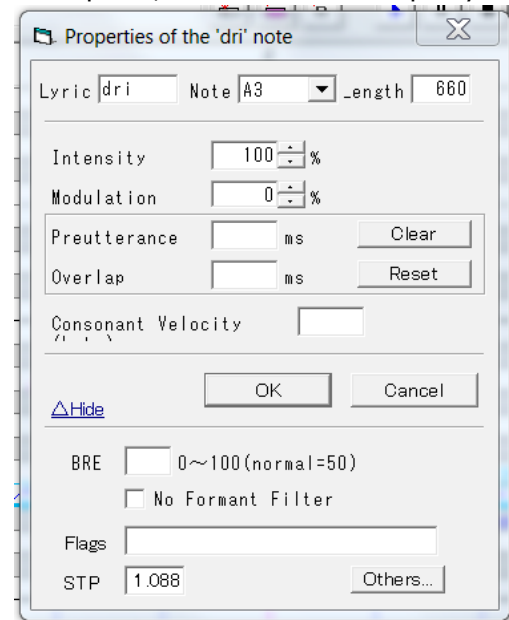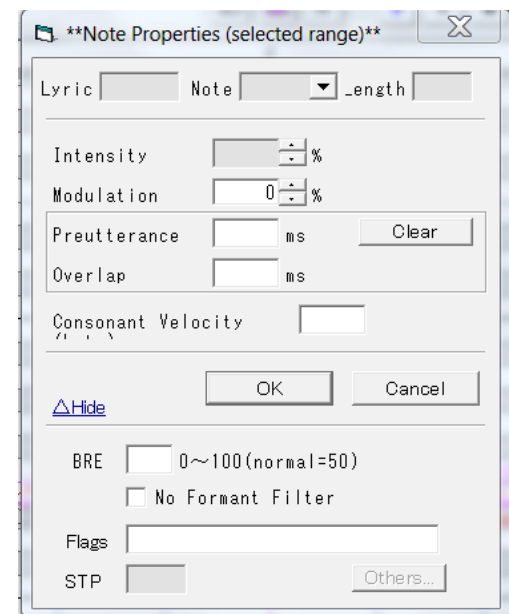

Illustration 13: Region properties dialog

12

breath, making it softer-sounding as well as quieter—at BRE100, the voice will become like a whisper, but may also sound rough and unpleasant. Using a low breathiness (less than 50) will take breath out of the voice, making it stronger and clearer.

   •   No Formant Filter: The formant filter adjusts the formant of the notes so that pitching is more natural. Leave this box unchecked for natural singing.

   •   Flags: The input field for flags, which control many different vocal parameters.

   •   STP: STP is set by OPT, ACPT, or by hand. It determines how much of the note is pushed behind the preutterance. This is done to keep longer consonants from crushing the previous note, so that all notes are audible.

Flags (Voice Parameters)

Flags control many aspects of the voice. You can find a fairly complete list of flags online on the UTAU wiki, on the UTAU User Manual page[6]. Some of the more useful flags are as follows:

Flag Name	Base Value	Setting Range	Description
g	0	-100 to +100	Formant shift. Positive g flags make the voice more mature/masculine; negative g flags make it more childish/feminine.
Y	100	0 to 100	Controls the "breathiness" of the consonant region. Low Y values (such as Y0) can make the voice enunciate more clearly, but they may also introduce noise.
H	0	0 to 100	A low-pass filter. Helps reduce noise in samples, but will also make them more muffled. Used with Y0 commonly.
P	86	0 to 100	Peaks compressor. Aligns the peak volume of the sounds prior to envelope editing, to help normalize volume.
B	50	0 to 100	The same as BRE, but a flag.
b	50	0 to 100	Breathiness (BRE) before the formant filter. Sometimes, this creates a clearer breathiness effect than BRE does.

Table data selected from UTAU Wiki[6]

Synthesis engines other than the provided "resampler.exe" may have different flag sets. See Appendix D: Sampling Engines for more information.

If you want to use a certain flag set over the entire file, enter it into the project properties dialog in the "rendering options" bar, shown in blue in illustration 14. In this example, I entered "b0" to my rendering options to change the default breathiness (pre-formant filter) to 0 for this file. I can still change the breathiness with flags, but 0 is now the default value rather than 50.
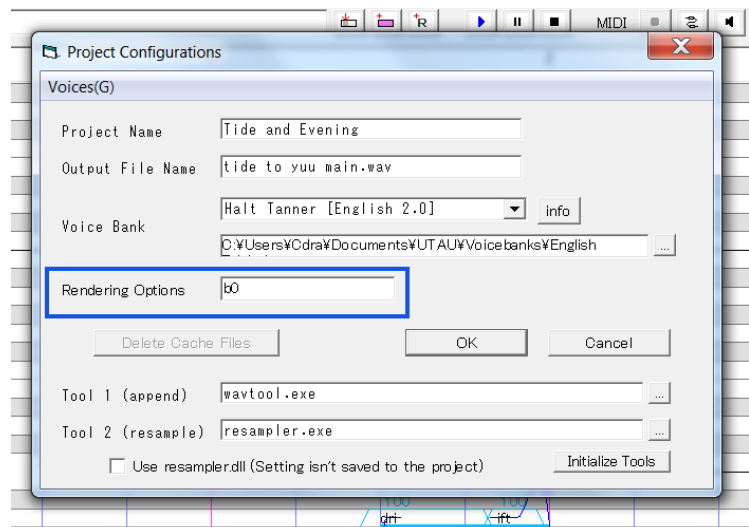

Illustration 14: Rendering options.

6http://utau.wikia.com/wiki/UTAU_User_Manual_-_7

Pitch Editing

Pitch editing is easily the most complex part of the tuning process. However, it is also the most powerful, as pitch edits are the best way to give the mechanical voice an emotional, human quality. There are three kinds of pitch edits: portamento, pitch-bends, and vibrato.

Portamento and pitch-bends are implemented using the same function, though they are very different. Portamento is the connectivity of the pitches between notes—how the pitches flow together. It keeps the pitch from jumping in a way that sounds choppy. Pitch-bends use additional portamento control points to help give the voice emotion by shifting the pitch within a note (ike how envelopes change the dynamics within a note). Vibrato is a more straightforward control, giving you the freedom to alter the vocalist's vibrato.

The first step to editing pitch is to open the "Pitch" dialog box. You can open this dialog for either a single note or an entire region; you can put vibrato, a specific portamento setting, or number of pitch control points on all the notes in your selected region. Note that rests will not be given control points or vibrato.

Illustration 15 shows the pitch control dialog. The portamento and vibrato boxes are both checked; you may uncheck them to remove all portamento (and pitch bending) or vibrato from the note.
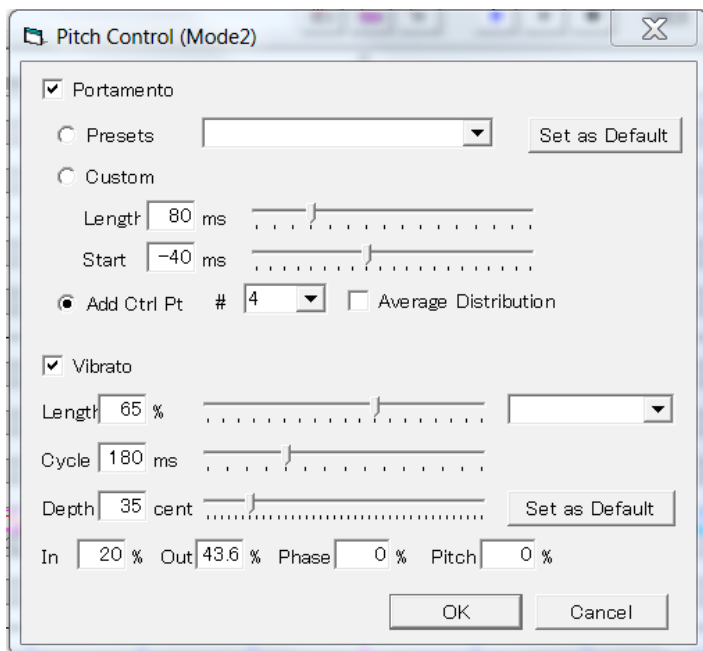

*Illustration 15: Pitch control dialog*

Portamento Options:
   •   Presets: This drop-down window contains various portamento styles you can use. You can click "Set as Default" to make your current settings the default portamento for all future notes.
   •   Custom: You can use this setting for custom portamento settings.
   •   Add Ctrl Point: Enter a number to add more control points, so that you can make custom pitch-bends and portamento shapes. Checking "average distribution" will distribute them through the note; otherwise, they will be gathered near the beginning of the note.

Vibrato Options:
   •   Length: The percent of the note that the vibrato occupies. Vibrato will fill the end of the note. The drop down window to the right has some vibrato preset options.
   •   Cycle: The period. Smaller "cycle" values result in faster vibrato.
   •   Depth: The amplitude, or how large the pitch variations are.
   •   In/Out: These values determine what percentage of the vibrato tapers in strength (from the beginning and end of the vibrato respectively).
   •   Phase: The phase shift; at what point in its cycle the vibrato starts.
   •   Pitch: This value indicates how much the vibrato is stretched up or down.
   •   Set as Default: Use this to set your current vibrato settings as the default for all future notes. This will make every note you create have that vibrato setting.

This dialog isn't exactly easy to use for pitch edits, but don't worry. You can edit pitches directly on the piano roll!

The first note in your selected region is the one you can edit the pitch on (see illustration 16). The little red blocks, called control points, move up and down when you drag them—except for the first and last points, which are stuck at the pitch of the previous and currently selected note respectively. Points must stay within the length of the note—that is to say, the first point can only go to the beginning of the previous note, and the last point can only go to the end of the current note. Aside from that, you can place the points anywhere within the note, and non-end points can move up and down any amount.
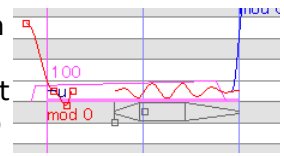
Illustration 16: Portamento.

Making a bend downward in the pitch between the first and second points is a common technique used to create strength in the voice; a similar bend up can create a strained sound. Experiment with different pitch-bend shapes to figure out what works for your song and vocalist.
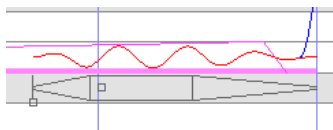

Illustration 17: Vibrato.

The gray box beneath the vibrato in your selected note is the vibrato controller, as shown in illustration 17.
- Drag the first bar to change the length of the vibrato.
- Click and drag the second bar to change how much the vibrato tapers in; the third to change how much it tapers out.
- Click and drag the top or bottom of the box make the vibrato larger or smaller.
- Click the inside of the box to drag and shift the vibrato up or down, so that it bends to a higher pitch or lower pitch slightly.
- The control point in the middle of the box adjusts the period of the vibrato (how fast the vibrations are), and the control point at the bottom left of the box adjusts the phase (where in the vibration cycle begins).

In general, vibrato is best placed on longer or higher notes, and should begin earlier in shorter notes. Like with pitch-bends, you should experiment with different vibrato shapes to figure oout what works.

There is no single way to tune the voice, so experiment with all of the tools at your disposal to get the effects you want!

Rendering

Once you've finished editing your ust, you should render it out for mixing. To do this, simply click "Project>Render wav file..." and select the location where you wish to save the render. I recommend that you do this twice in order to optimize your crossfade. Save the file then render it. Then, select the entire ust and click "OPT" in the top right-hand corner of the window. Once it finishes optimizing and shows "Cache of selection is removed, OK?" click "OK" to complete the optimization process. Then render your file again the same way as before; the second render will be clearer than the first.

As a note, if you wish to only render a small section of a track (such as for a demo), you can select that section, play it (with the play button), then click "Play>Save Last Play..." to save your most recent playback as a wav.

Final Remarks

Learning to use UTAU is a rewarding experience, and it can be very useful in the production of

music. It's also a lot of fun to experiment with UTAU! Like any instrument, however, it takes practice to become skilled at creating electronic vocals. This guide is intended as an introduction to the concepts of the program so that you can begin the learning process with the basics in hand.

Of course, it's entirely possible that you will run into technical issues with the program—if you do, I recommend checking out the Overseas UTAU Forum[7], which has lots of threads dedicated to helping people use the program.

Most of the information in this guide was drawn from my own experience with the program. As such, there is a chance that some of my information is inaccurate. If you've read this guide and found that something I said was not correct, please contact me at cdra1617@gmail.com so that I can correct the guide for future users!

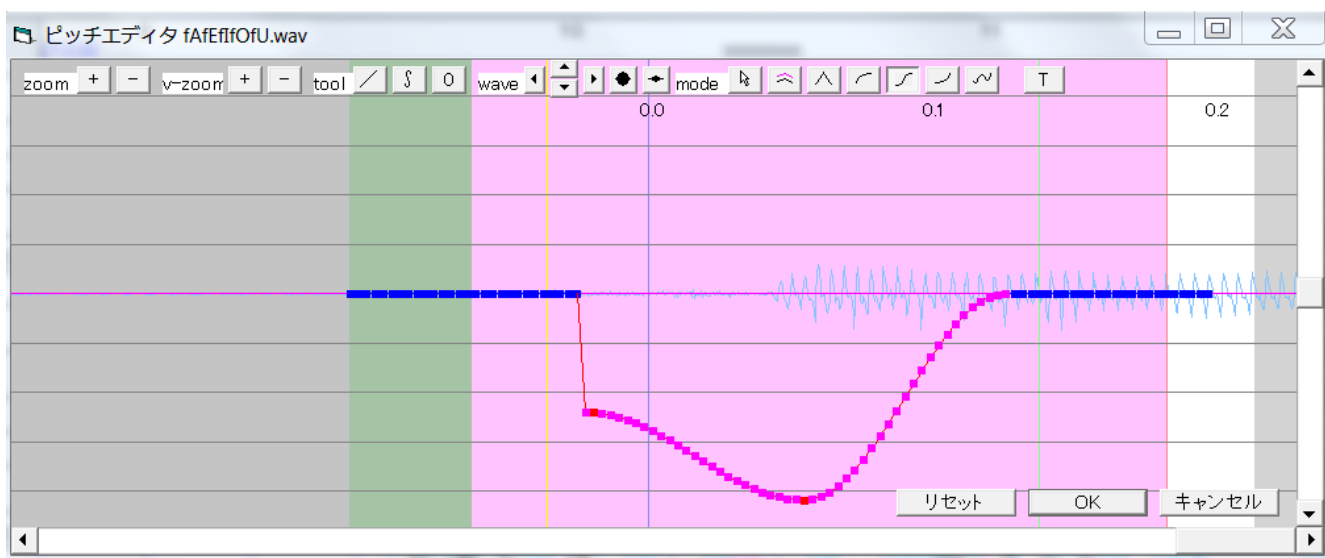## Appendix A: Mode1 vs Mode2 Pitch Editing

Illustration 18: The mode1 pitch editing window.

Mode1 pitch editing was the only method of editing pitch in UTAU when it was first released. However, mode2 is much more functional and much easier to use. Mode1 is based on drawing pitches into a separate editing window, as shown in illustration 18. Mode1 does not have any form of automatic portamento (and creating portamento manually is extremely difficult), and vibrato must be drawn by hand. Because of this, I do not recommend that you use mode1 editing at all.

If you download a premade ust and find that the pitch editing is in mode1, first make sure that the preutterance parameters are locked for the ust (as the pitch edits will move if the parameters are changed), then click the "Mode2" button in the icon palette. This will convert the ust to mode2; however, the pitch edits will be lost. You should then click "Trace" to see the pitch edits as gray lines, which you may then trace with mode2 pitchbends in order to keep the pitch edits. Alternately, you can simply convert the file to mode2 and rebuilt the pitch edits from scratch. Either way works.

---

7   http://utaforum.net/index.php

Appendix B: CV VC Editing

CV VC is the youngest recording style used in UTAU, but it has been growing in popularity recently. Editing usts for CV VC is a bit different than editing them for any other recording method, as CV to/from VCV (Japanese) conversions are easily done with a plugin. I recommend shinami's plugin tutorial[8] for information on where to find these plugins, as well as how to install them.
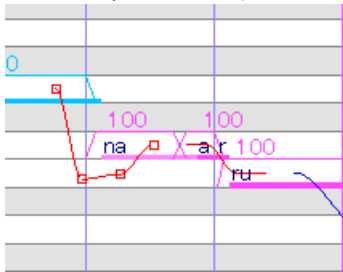

Illustration 19: CV VC Japanese

To convert a ust to CV VC, you split a small note off the end of each note (using Ctrl+Shift+Drag) and enter the "VC" lyric to that note. For Japanese, this is especially simple; just enter the same vowel as the previous note (the "a" in "na" for illustration 19) and the same consonant as the second note (the "r" in "ru" here). This creates a "manual VCV" blend, blending notes only along the same sound rather than between different ones.

For non-Japanese CV VC voicebanks, such as English, the lyric you enter in the new, small note is different. You will put the phoneme that completes your word in the second note: recall how in the word "adrift," the small ending note had the "ift" sound in it.

When the note you are splitting is followed by a rest, you add the VC after the note as shown in illustration 20 (the "Iv" note). Instead of splitting the VC part of the note out of the main note, you split it out of the rest that follows it. This is done to keep the singing on-time. You may also include a trailing-out vowel sound in the same way, such as "i-" or "i -;" sometimes these are included in CV or VCV voicebanks as well.
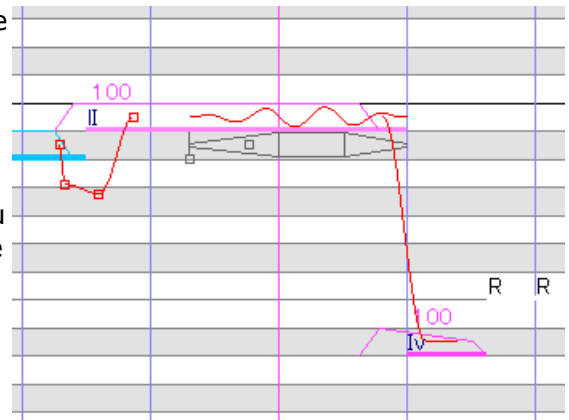

Illustration 20: VC where the next note is a rest

Appendix C: Importing MIDIs and VSQs

Importing a MIDI or VSQ file to UTAU is a simple process, but it is rather error-prone due to bugs in the program. To import a file, simply "File>Import..." and use the resulting file browser dialog to find the file you wish to import. If you do not see the file, make sure that the drop down window in the bottom right of the dialog (right above "Open" and "Cancel") displays the type of file you wish to import. If it says "SMF Format File," it is trying to import a MIDI, and if it says "VSQ Format File," it is trying to import a VSQ.

After selecting the file you wish to import, click on the track with the vocals. You will have now imported your file to UTAU. If you're lucky, it will have imported without any problems; however, in many cases this will not be the case. Often UTAU will import VSQ files slightly off-time, such that each note is just a tiny amount too short; with MIDIs, it is prone to not import at all due to issues with the MIDI's format. Make sure to double-check your MIDI or VSQ after importing it to be sure that there are no problems.

---

8   http://shinamieba.deviantart.com/art/UTAU-Plugins-Tutorial-270068751

Appendix D: Sampling Engines

One of the great things about UTAU is the ability to change what sampling engine you're using. To explain better, the sampling engine is constructs the singing from the voice samples, pitching them to the specified locations on the piano roll and stretching them to the correct lengths. The sampling engine that comes with UTAU is "resampler.exe," but there are several others developed by both Ameya and third-party users.

The three most used sampling engines are resampler.exe, fresamp.exe, and TIPS.exe.

- Resampler.exe: Resampler is known to maintain the strength of the voice well, and so is often the engine of choice for stronger voices. However, it also reacts especially poorly to lower-quality voicebanks, and may cause extra buzzing in those banks. The newest resampler handles breathiness extremely well compared to other engines.
- Fresamp.exe: Fresamp is also good for strong voices, but it makes voices incredibly nasally without the use of the F flag. It is often considered to be a bit clearer than resampler.
- TIPS.exe: TIPS tends to behave well with soft banks and low notes, but creates a distinct noise on some samples. It has also been known to glitch on some voicebanks. Rather than using the .frq files that come with voicebanks and are used by resampler/fresamp, TIPS generates its own ".pmk" pitch map files. Only the H, P, t, and g flags work with TIPS. It also has its own flag called R, which causes the sampler to regenerate .pmk files.

You can find download links to each of these engines, as well as links and information about more engines, on UTAforum in shinami's Resampler Directory[9].

Different engines work well with different voicebanks. Often, the optimum engine is listed in the voicebank's readme file, but experimenting with different engines is encouraged to find the one that best suits your voice and song.

You can change the sampling engine in the Project Properties dialog; it is at the bottom of the dialog under "Tool 2 (resample)". Simply click the "..." to the right of the input bar to browse for the sampling engine you want to use.

---

9   http://utaforum.net/index.php?topic=550.0