# C+V ENGLISH CREATION GUIDE

**Version 103**



**What is C+V?**
C+V is an UTAU format where you record consonants and vowels separately and then string them together in UTAU.

**Why C+V ENG?**

- Very short production time - small reclist, short oto.
- Very beginner friendly - short recordings means more time to re-record if there's a pronunciation mistake or other technical issues. Can also be used as a 'warm-up' for longer, more 'proper' ENG recordings (Especially ARPAsing)
- Easier than ever to use - OpenUTAU makes using C+V ENG very easy with phonemizers - you just put a word in and OU automatically assigns the phonemes for it a la commercial synths like SynthV, unlike original UTAU where you need to put in the phonemes one by one.

If you are interested in how C+V ENG sounds or how you use it, Veria, my C+V ENG voicebank, is available for download and you can hear her demo reel.

This tutorial will go through the creation of a C+V ENG bank from start to finish, as well as usage/tuning tips and advanced tips.

**Things to Download**

While this tutorial is geared towards beginners, I am assuming that you have used or at the very least downloaded and read the 'Getting Started' page on these apps:

- OpenUTAU
- VLabeler
- Recstar(recommended)/Oremo

I would also assume you have a DAW (Digital Audio Workspace) available.

# 1. RECORDING

Veria's recording list is available to download on her website. This recording list, and this tutorial as a whole, uses ARPABET. Once you download the reclist, you can download Recstar (recommended) and follow the tutorial on Github or Oremo. To make sense of ARPABET, you can either listen to Veria or Biggity's recordings, go to the Phoneme List section or the SynthesizerV phoneme reference.

**FOR PEOPLE WHO ARE PLANNING TO DO A MULTIPITCH VOICEBANK:**
Please jump to the Multipitch section before recording. There are aspects during recording and otoing that will deviate from these two sections, and a dedicated toolkit will be available for you.

Every line in the recording list is said continuously, even the ones that have multiple phonemes.
EX: "aa-r" is recorded as "aarr" (like a pirate) instead of "aa(gap)er". "b-ax-b" is recorded as "baaab" (like bub) instead of "bee ax bee".

Your recordings should:

1. Have silence in the beginning and end, and;
2. Be held for around a second long (not including silence).

For consonant recordings, you are able to change the middle vowel to your liking. Even when recording Veria I switched vowels multiple times before settling on [ax]!

About differences between Veria and Biggity's recordings

I (Chevrefee) am an ESL (English as a Second Language) while Biggity is a native English speaker. There are some differences in the recording samples that you will find (most notably [ae]), but both pronunciations are valid and will work - it is just a matter of accents. This guide will follow Veria's accent.

# 2. OTO

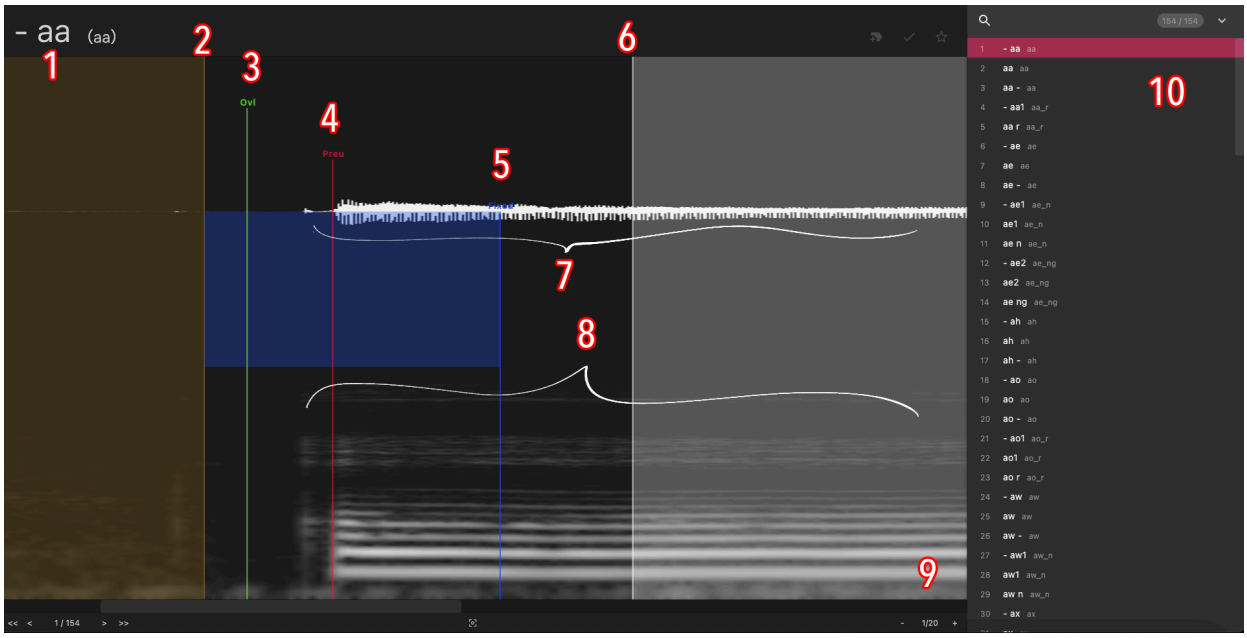
(This guide exclusively uses VLabeler and will not cover SetParam)

After recording, you will need to oto your recordings. If you used Veria's recording list, you can also download Veria's base oto as a starting point.

This oto guide wouldn't be possible without techniques and descriptions adapted from various otoing tutorials combined with user feedback - Full credits on the credits page.

## Intro To VLabeler/Otoing

Put the base oto in the same folder as your recorded samples. After loading up the oto file in VLabeler, you will be greeted with some coloured lines on the main window, and the name on top, called the alias. These lines will determine how your voicebank will sound in UTAU.




1. Alias and (sample) - Name of the sound and (name of the recording)
2. Offset - Start of the sound
3. Overlap - Where the notes crossfade onto one another
4. Preutterance - Where the note starts
5. Fixed - Which part doesn't get looped/stretched
6. Cutoff - End of sound
7. Waveform - Visualization of sound
8. Spectrogram - Visualization of sound frequency
9. Resolution - Zoom in/out of the viewport. Recommend to set it to 1/10 when otoing
10. Entry List - List of all aliases and its sample origin

(Based on Yin-P's otoing guide)

There are three types of aliases in this oto - The starting alias [- C] , middle alias [C], and ending alias [C -]. It corresponds to the start, middle, and end of your phoneme.

There are 2 major categories of phonemes, Consonants and Vowels, and these categories are split into smaller categories.

Vowels are split into regular vowels and diphthongs, while Consonants are split into hard consonants, soft voiced consonants, soft unvoiced consonants,

semivowels, and special consonants. Each category will have different ways to oto their aliases.

**Please pay attention to the placement of the lines on the waveform and the spectrogram in the visual aids provided. Drag the Preutterance line first, and then adjust the rest.**

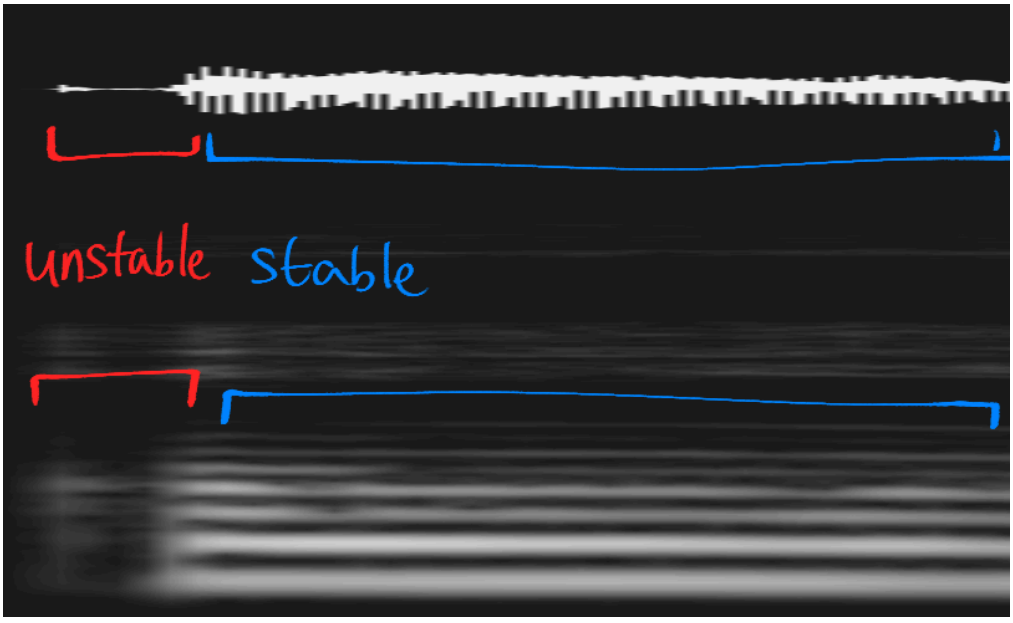
**Right clicking on a section will play the section's audio, and space will play the audio from offset to cutoff.**

## Terms FAQ/Before You Oto

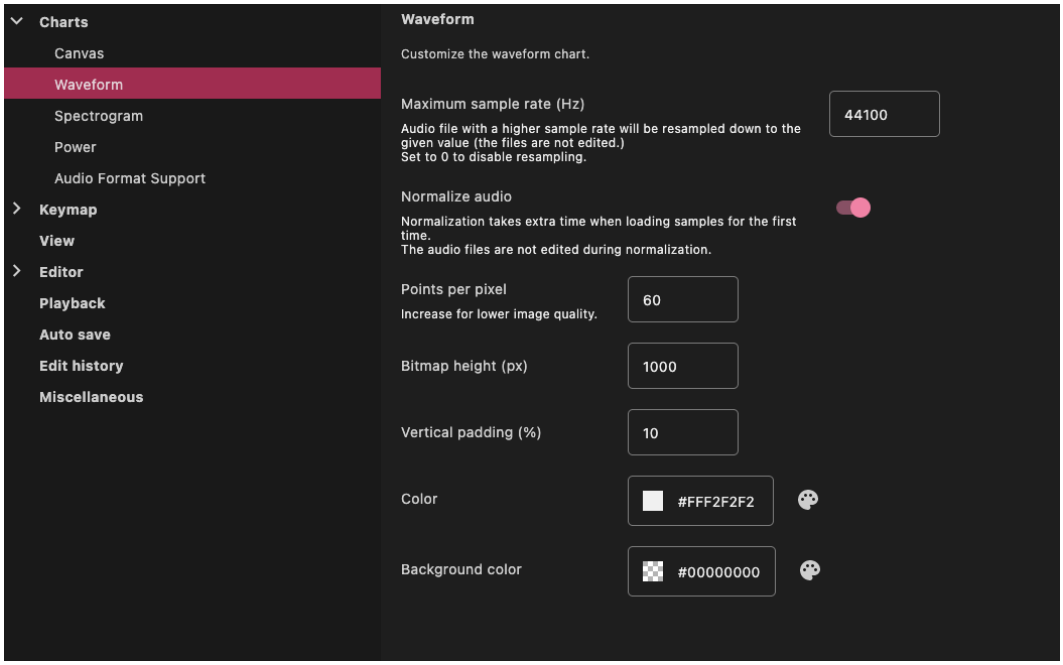
**What does ⅓, half or twice of preutterance/fixed mean?**
To put it simply, that's where you put the overlap line on the section between the preutterance and the offset (or fixed and the previous line). If it's a fraction, it goes left. If it's a multiplication, go right.

EX:
Overlap is ⅓ of preutterance. Set up the preu. and offset, and slide the overlap to one-third of the gap between the two lines.

**Vowel/Consonant stable? Solid spectrogram?**
Look at the waveform and spectrogram. If the waveform is around the same height, or the spectrogram has clear, solid lines, that means the vowel/consonant is stable.



If you are having difficulty in gauging what's stable and what's not, I would recommend going to Settings -> Preferences -> Waveform and enabling Normalize audio.




**I think the starting part of my consonant sample is too long to apply the oto-ing techniques.**
You would need to edit out a chunk of the starting part. This usually happens in hard consonants when you slur the recording (EX: recording g_ax_g as “nggug”). Silence the parts where it's not a pure consonant with a DAW.

**Is there a way to test out the oto in real time?**
Sort of! If you open both VLabeler from OU (Tools -> Singer -> Edit in VLabeler) and save in VL after you modified the oto, OU should update the oto automatically. **Be warned that if you have a premade ustx and you play it without editing the base oto first, OU might crash!** I would suggest using a fresh project and input words as you oto if you want to test it out while otoing. If it stops syncing, go to Tools -> Clear Cache.

**I would oto this differently.**
That’s totally okay - These techniques are just the ones that work for Veria.

**I followed the previous oto tutorial - Would it still work well when it’s pretty different from this version?**

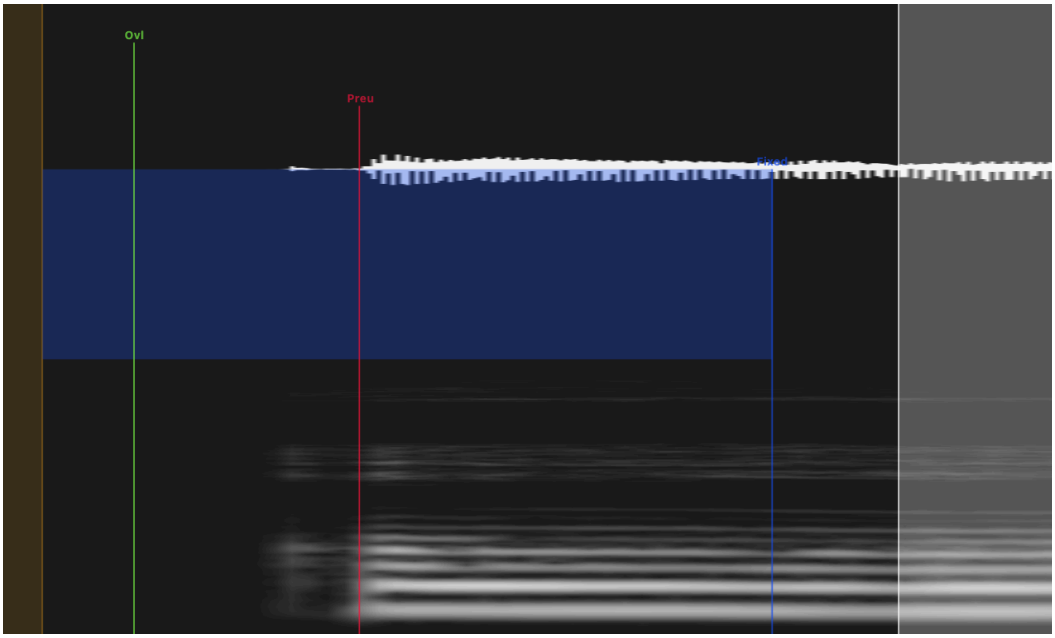
It should still hold up, though I recommend following these steps instead as it's most compatible with the base oto and phonemizer and thus will be smoother.

## a. Vowels

### i. Regular Vowels (aa, ae, ah, ao, ax, uh, uw, ih, iy, eh, er, ur)

[- V] Starting

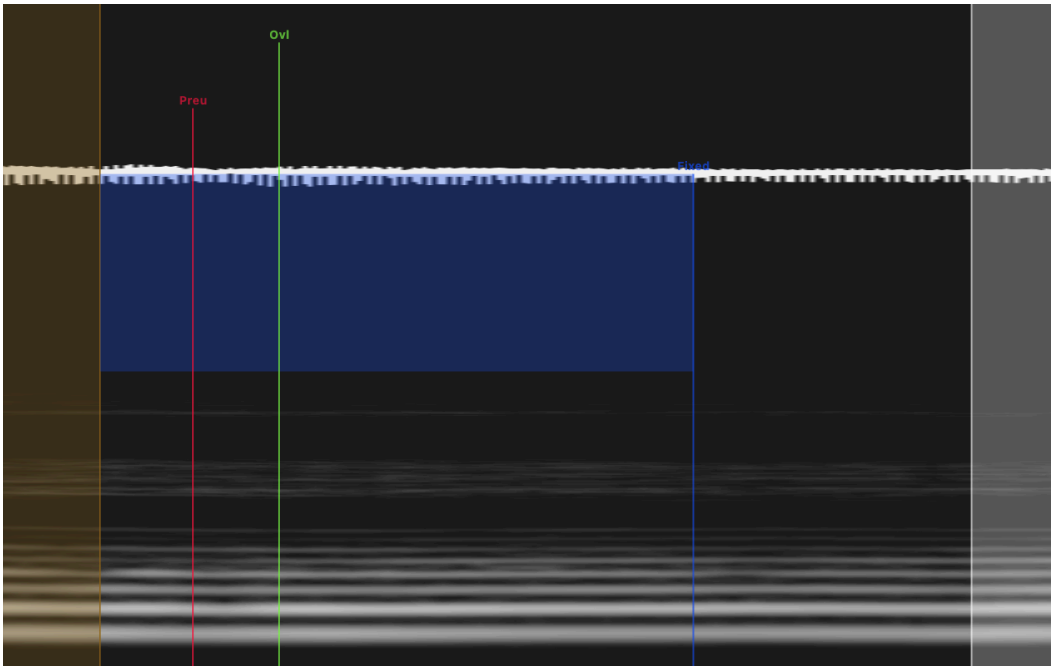
**Preutterance**: Start of vowel (Spectrogram goes solid)
Offset: Silence before vowel
Overlap: ⅓ of preutterance
Fixed: When the vowel gets stable in pitch/volume
Cutoff: When the vowel starts to become unstable/waveform increases/decreases in height. If vowel is stable throughout, put it at around 1.5 times of fixed




[V] Middle

Note: oto like japanese vowels
**Preutterance**: When vowel is stable (Spectrogram is solid)
Offset: When vowel is stable (Place the preu. so that the offset is also on stable part)
Overlap: Twice the preutterance
Fixed: When the vowel gets stable in pitch/volume

Cutoff: When the vowel starts to become unstable/waveform increases/decreases in height. If vowel is stable throughout, put it at 1.5 to 2 times of fixed




[V -] Ending

Note: oto like japanese ends ([v] R)

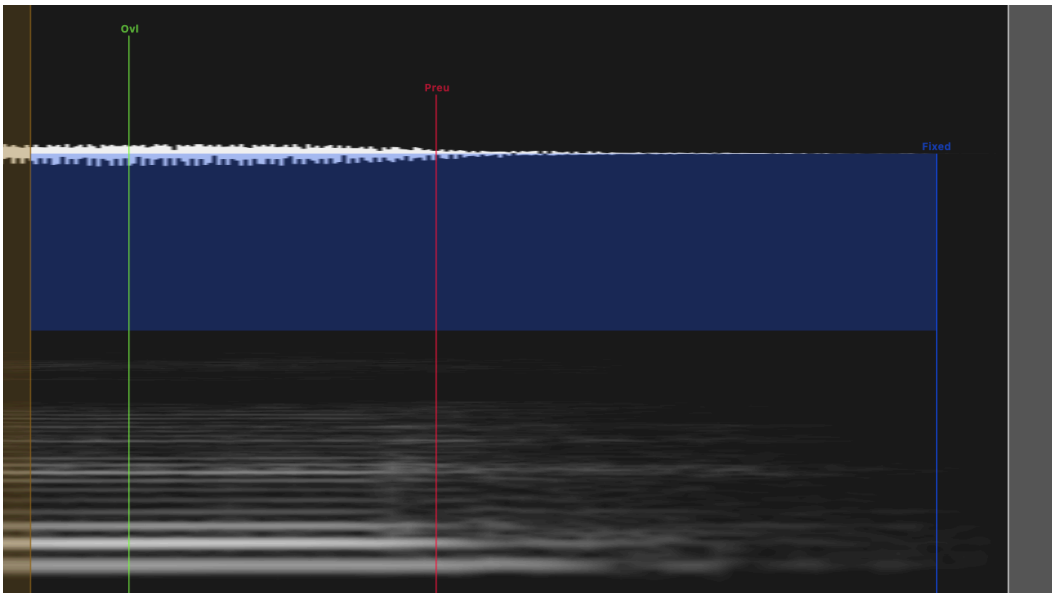
**Preutterance**: Between the vowel and start of fadeout/ending breath (When spectrogram starts fading off)

Offset: When the vowel is still stable (Spectrogram is still solid)

Overlap: ½ of preutterance

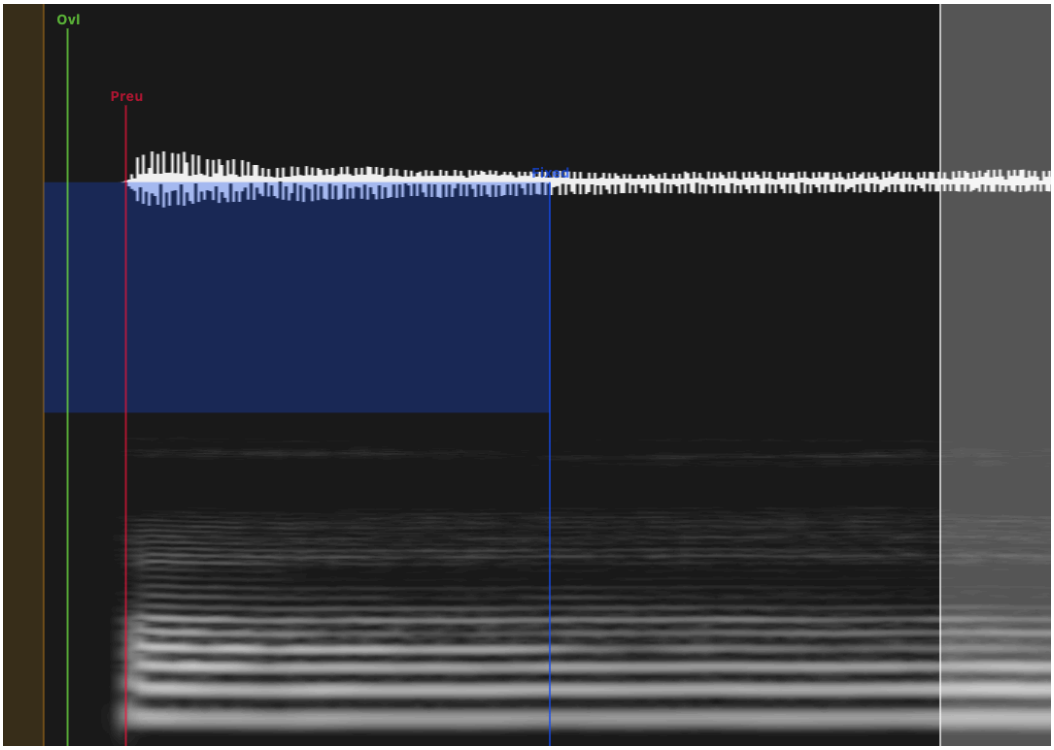
Fixed: Tail end of waveform

Cutoff: A chunk of silence after fixed. If there's unwanted additional sounds, leave it out (leave it on the right side of the cutoff)

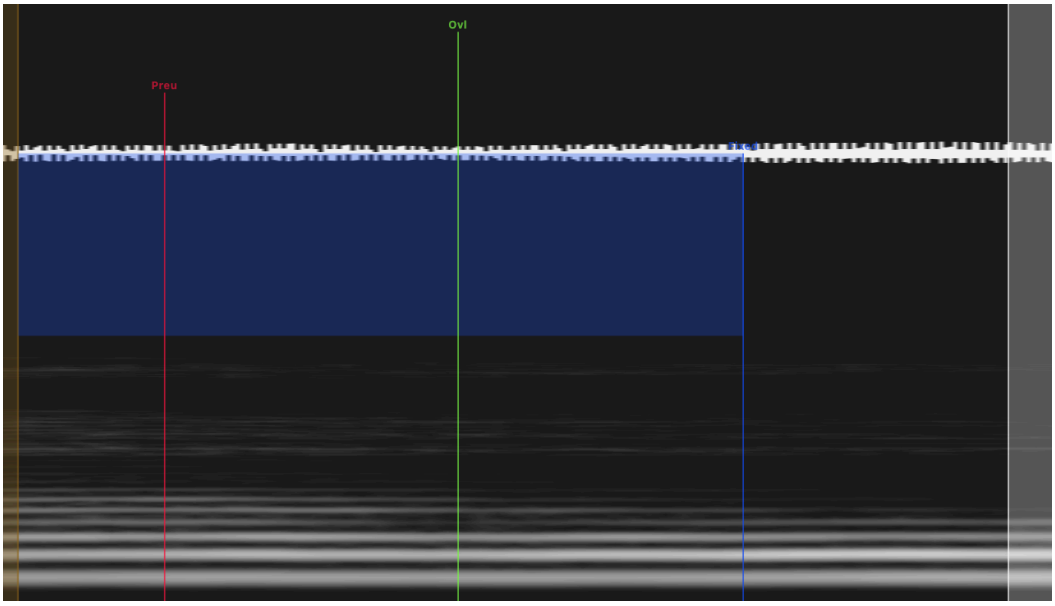



ii. Diphthongs (aw, ay, ey, ow, oy)

[- V] Starting (UPDATE)
**Preutterance**: Start of vowel (Spectrogram goes solid)
Offset: Silence before vowel
Overlap: ⅓ of preutterance
Fixed: When the vowel gets stable in pitch/volume
Cutoff: When the vowel starts to become unstable/waveform increases/decreases in height. If vowel is stable throughout, put it at 1.5 to 2 times of fixed




[V] Middle (UPDATE)
**Preutterance**: When first vowel is still stable
Offset: When the first vowel is stable (Place the preu. so that the offset is also on stable part)
Overlap: Twice of preutterance
Fixed: When the vowel gets stable in pitch/volume
Cutoff: Before second vowel starts

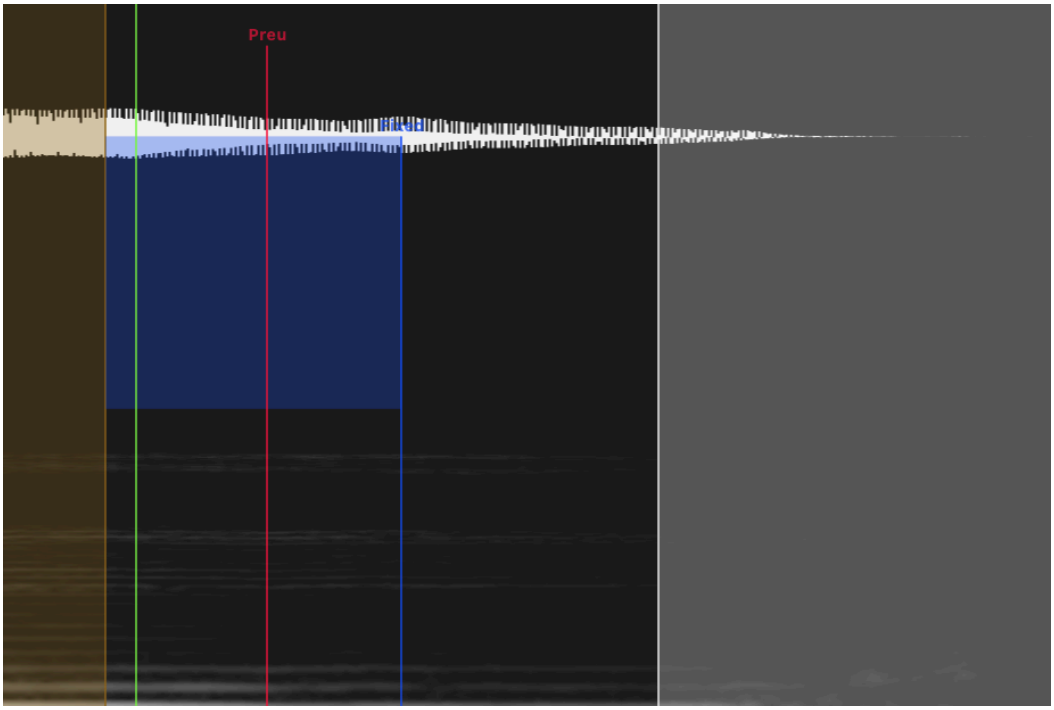



[V-] Second Vowel (NEW)
This entry is adapted from Astel's C+V English format. Because of the way the phonemizer works, this entry is somewhat optional, but for a very noticeable boost in smoothness you probably should configure it.

Note that [V-] and [V -] are NOT the same thing. As of 14/10/24, this entry is included in the base oto.

**Preutterance:** Between first and second vowel
Offset: When first vowel is still stable
Overlap: ⅓ of the preutterance
Fixed: When the second vowel gets stable in pitch/volume

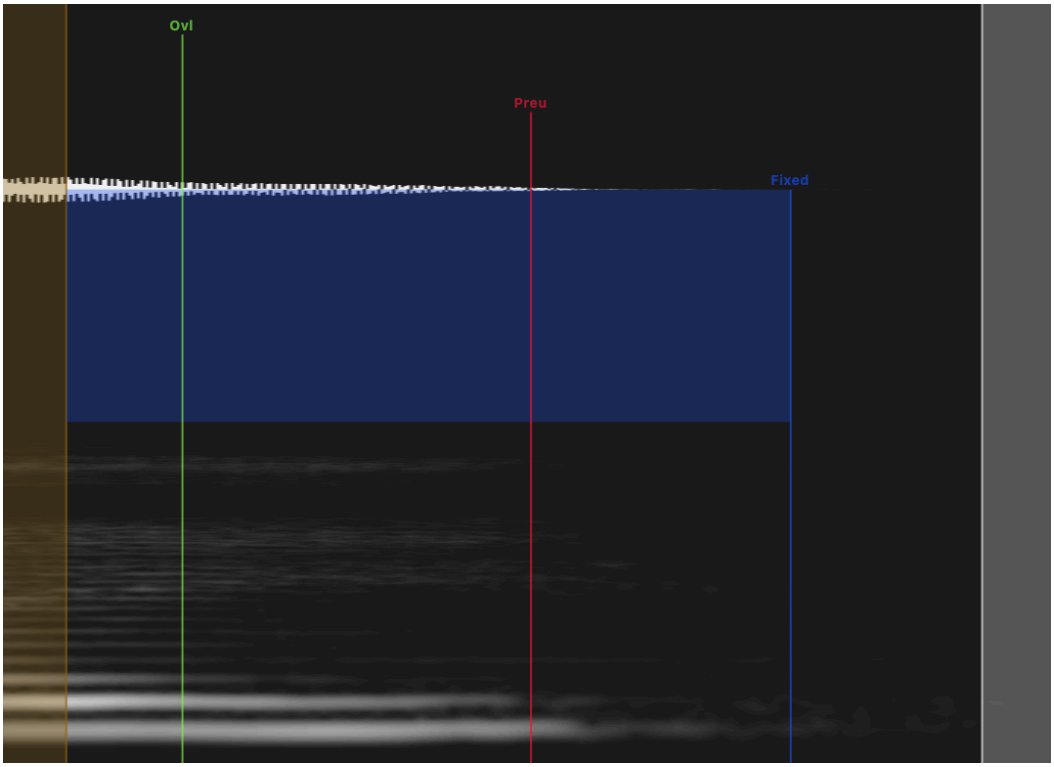
Offset: ¾ of gap between start of second vowel and the tail end of the waveform




If you would like to test these [V-]s, I would recommend typing these into OU's piano roll and adjusting your preuttenance if it sounds choppy/disconnected from the [V]s:
Play Pow Clone Vie Ploy Play

[V -] Ending
**Preutterance**: Between the second vowel and start of fadeout/ending breath (When spectrogram starts fading off)
Offset: When the second vowel is stable in pitch/volume
Overlap: When the second vowel is fully detached from the first vowel*
Fixed: Tail end of waveform
Cutoff: A chunk of silence after fixed. If there's unwanted additional sounds, leave it out

* EX: It's hard to describe, but if you are for example otoing [ey], the overlap should be when it starts to sound purely like the [y] sound. This is also when the spectrogram starts to go slightly less solid

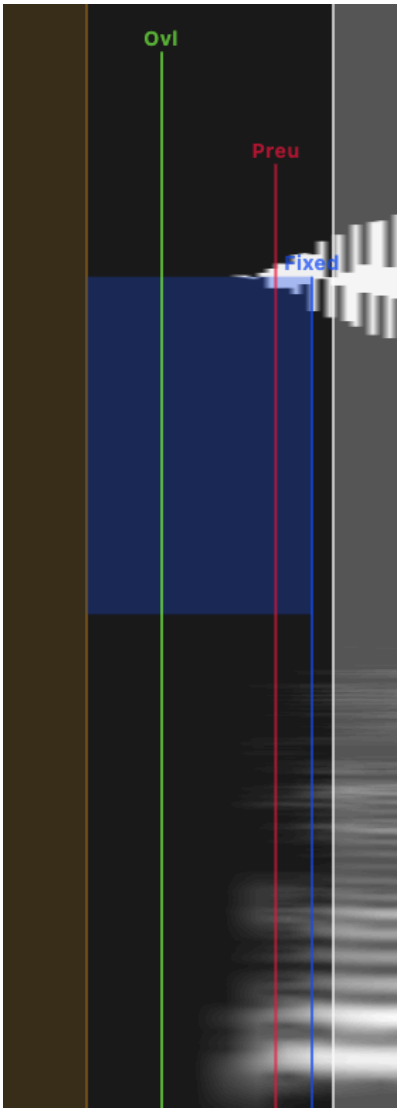



## b. Consonants

i. Hard Consonants (b, g, ch, d, p, t, k, jh, dx) and Special Consonants (q)

[- C] Starting

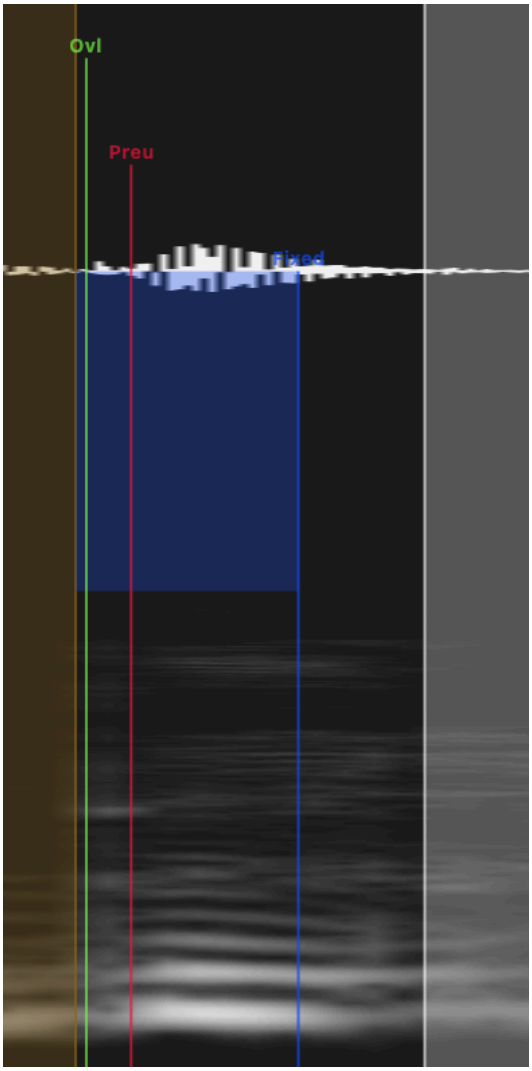
**Preutterance**: Start of consonant (Spectrogram goes solid)
Offset: Silence before consonant
Overlap: ⅓ of preutterance
Fixed: On the end of the consonant sound. Get as much consonant as you can, but leave a small gap from cutoff
Cutoff: Before the vowel starts




[C] Middle
Veria's base oto puts most of these entries on the ending part of the sample because it's more pronounced than the starting part. Feel free to slide the lines to the front/back, whichever sounds more pronounced.

**Preutterance:** Start of consonant
Offset: Silence/Gap between vowel and second consonant
Overlap: ⅓ to half of preutterance. Depends on your recordings. If ⅓ doesn't feel right, slide it to half
Fixed: On the end of the consonant sound. Get as much consonant as you can, but leave a small gap from cutoff

Cutoff: Before the consonant fades out (before ending breath fades in)




[C -] Ending

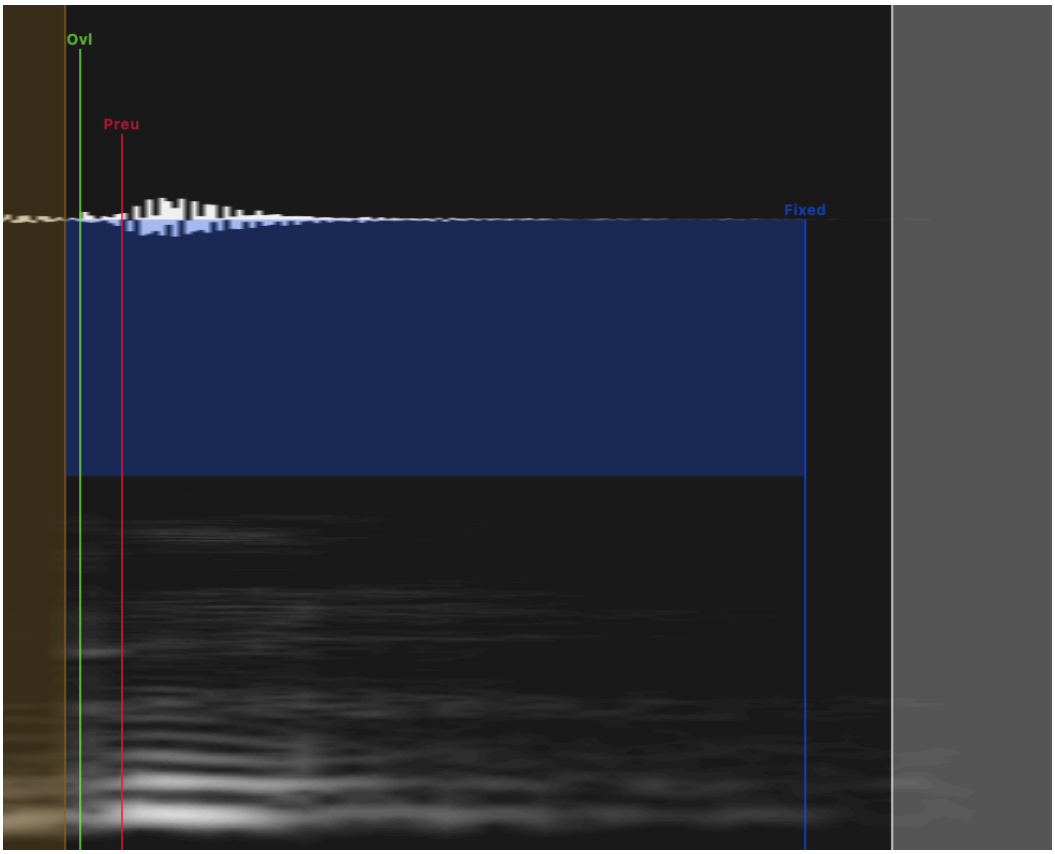
**Preutterance:** Start of second consonant

Offset: Silence/Gap between vowel and second consonant. This should be around/closer to the second consonant than [C] middle

Overlap: ⅓ to half of preutterance. Depends on your recordings. If ⅓ doesn't feel right, slide it to half

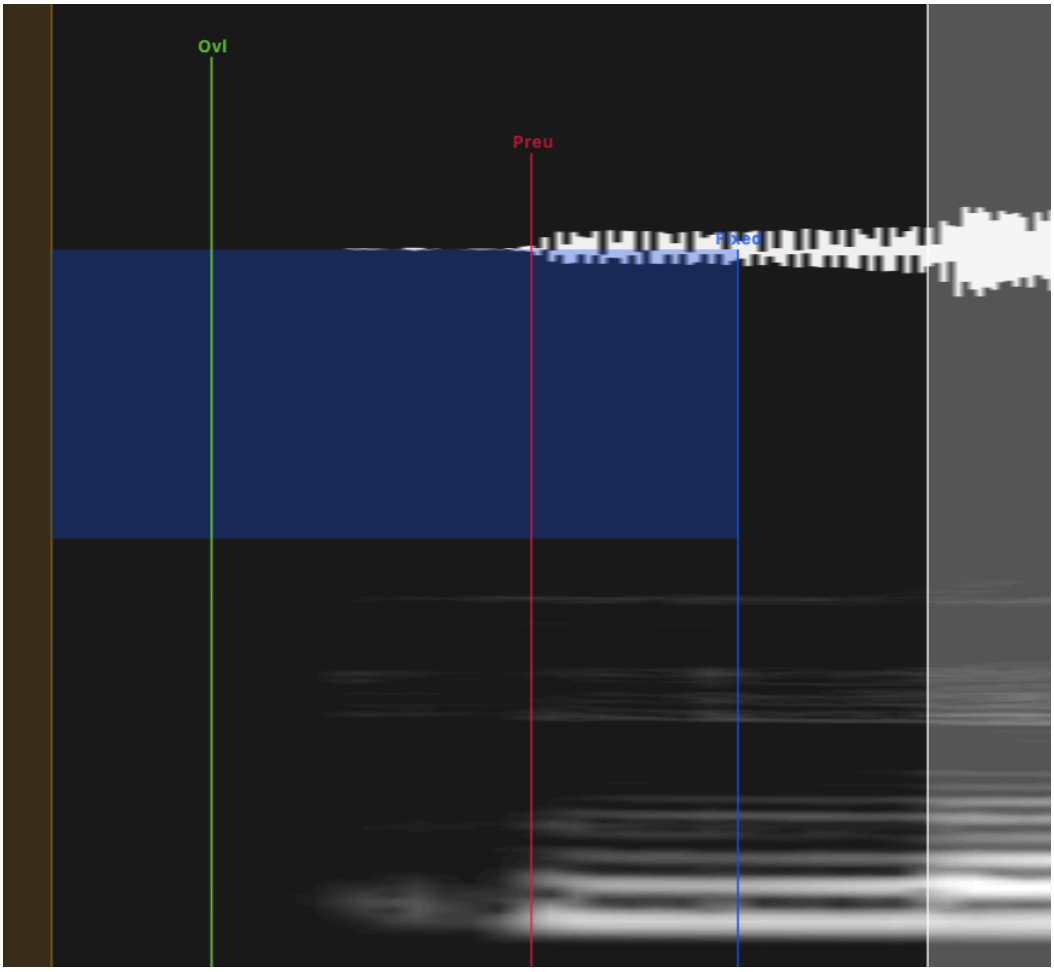
Fixed: Tail end of waveform

Cutoff: A chunk of silence after fixed. If there's unwanted additional sounds, leave it out




ii. Soft Voiced Consonants (l, m , n, ng)

[- C] Starting
**Preutterance:** Start of consonant
Offset: Silence before consonant
Overlap: ⅓ of preutterance
Fixed: When consonant is stable
Cutoff: Before the vowel starts




[C] Middle

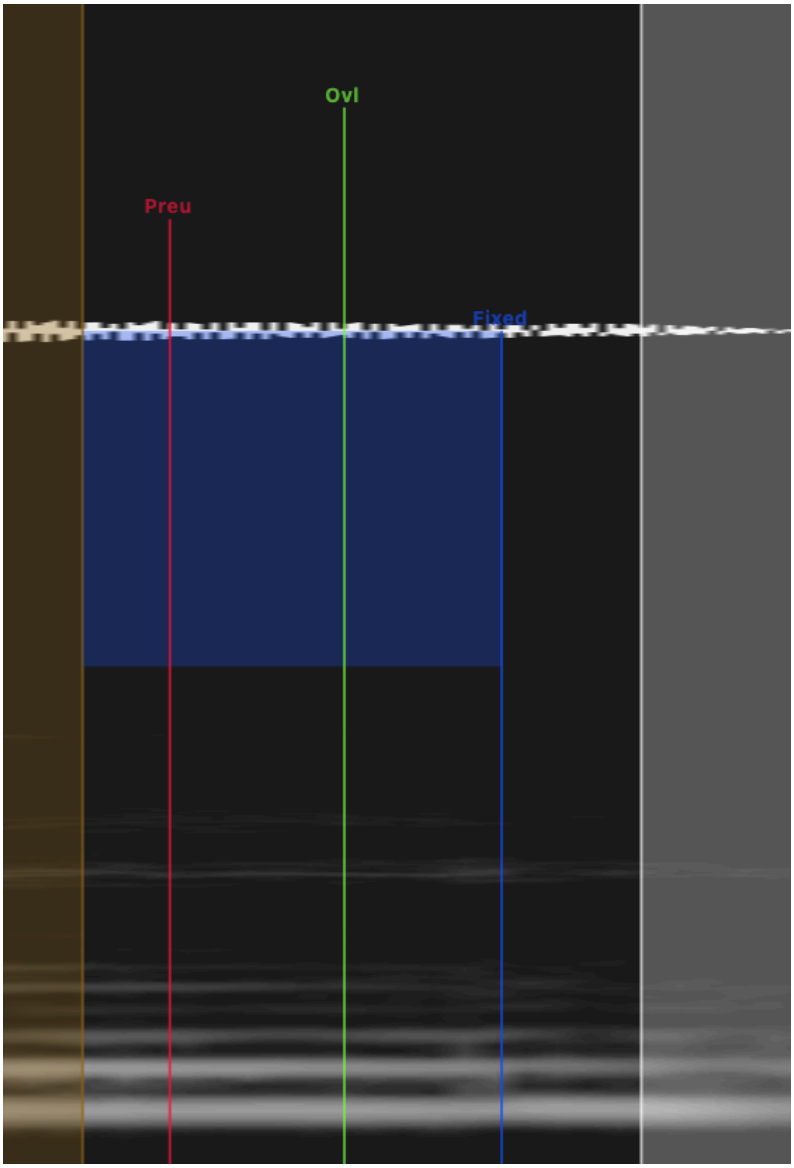
**Preutterance**: When the second consonant is stable

Offset: When consonant is stable (Place the preu. so that the offset is also on stable part)

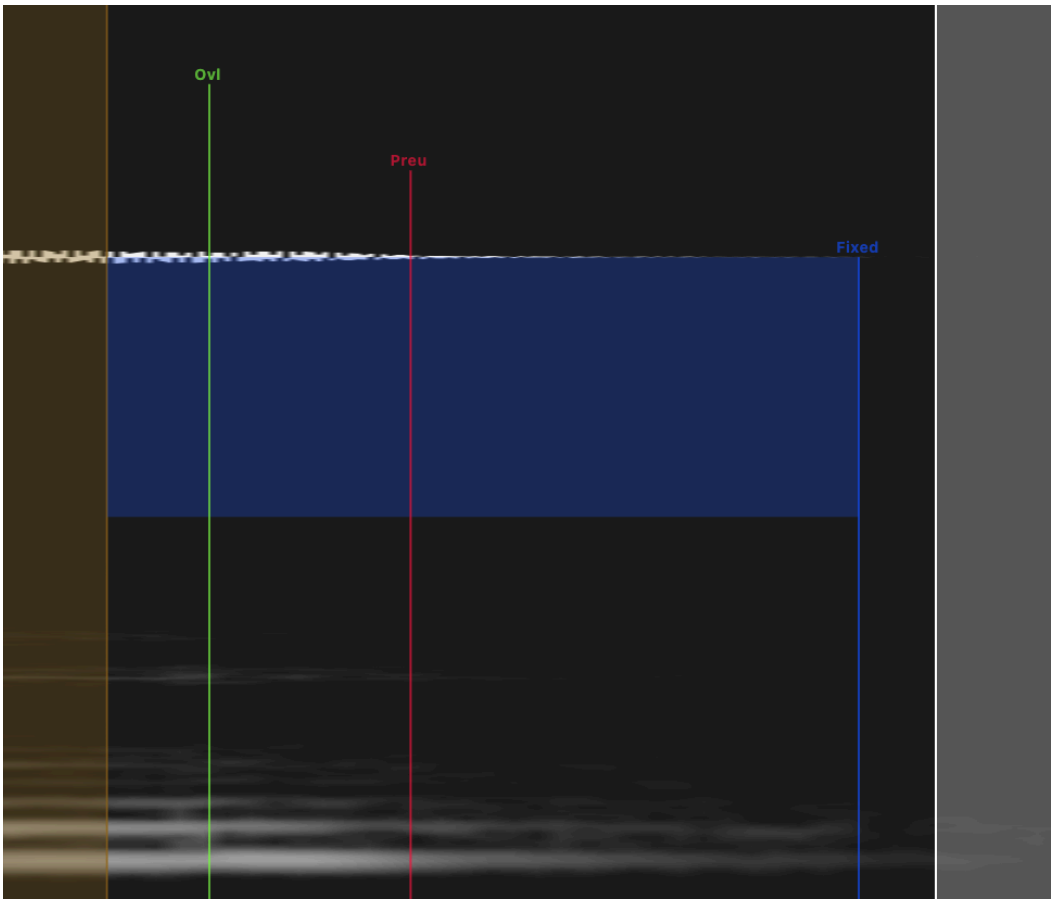
Overlap: Twice the preutterance

Fixed: When the consonant gets stable in pitch/volume

Cutoff: Before the consonant fades out (before ending breath fades in)

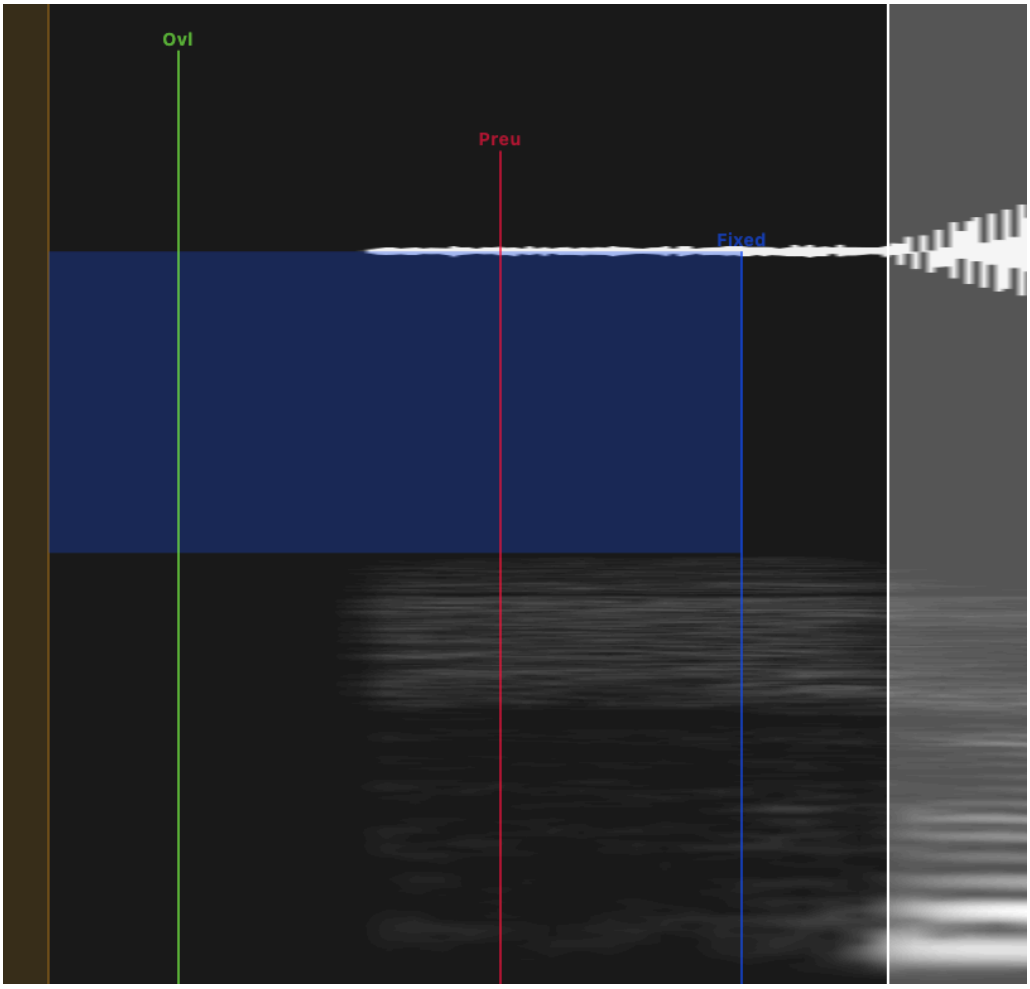



[C -] Ending
**Preutterance**: Between the consonant and start of fadeout/ending breath
Offset: When the second consonant is still stable
Overlap: ⅓ to half of preutterance
Fixed: Tail end of waveform
Cutoff: A chunk of silence after fixed




iii. Soft Unvoiced Consonants (dh, f, hh, s, sh, th, v, z, zh)
These entries might be a little harder due to the spectrogram not being solid. Pay closer attention to the waveform.

[- C] Starting
**Preutterance:** Start of consonant
Offset: Silence before consonant
Overlap: ⅓ of preutterance
Fixed: When consonant is stable
Cutoff: Before the vowel starts




[C] Middle

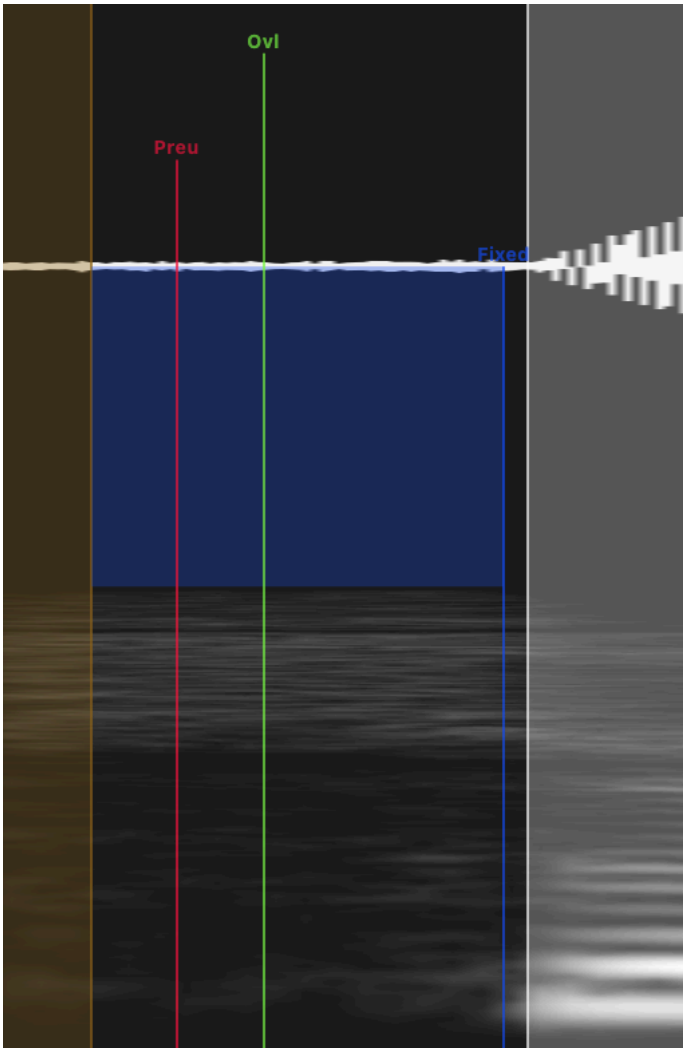
**Preutterance:** When consonant is stable

Offset: When consonant is stable (Place the preu. so that the offset is also on stable part)

Overlap: Twice the preutterance

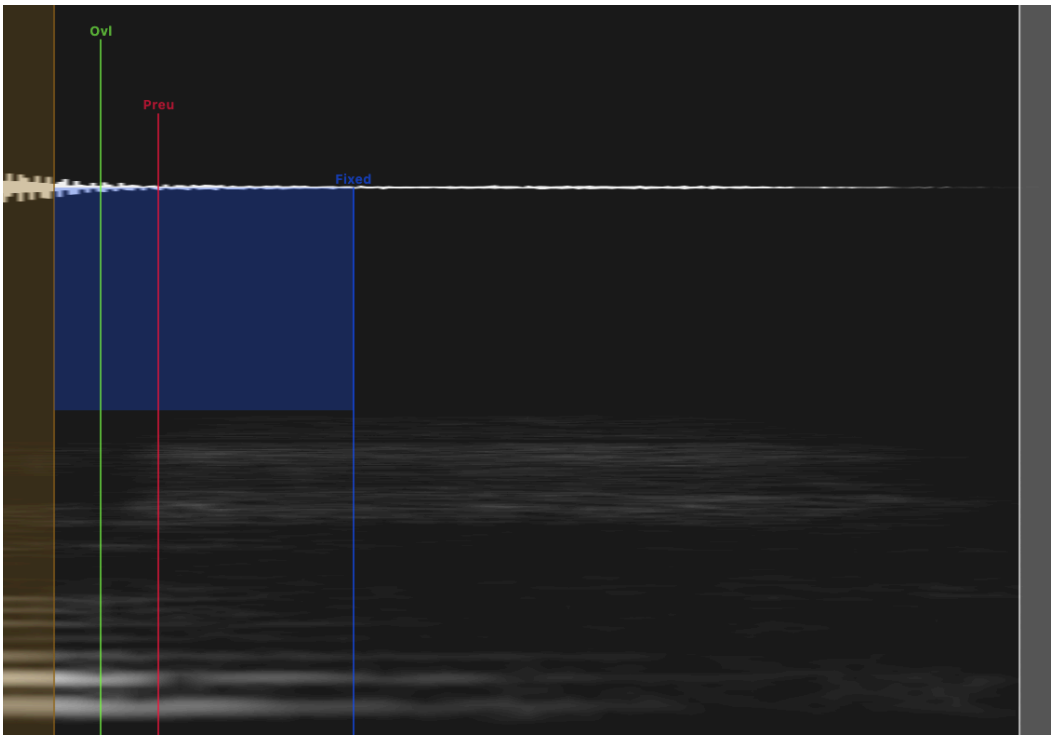
Fixed: On the end of the consonant sound. Get as much consonant as you can, but leave a small gap between fixed and cutoff

Cutoff: Before the vowel starts or before the consonant fades out (before ending breath fades in)




[C -] Ending (Adapted from Cytalai's Input)
**Preutterance**: At the start of consonant fadeout
Offset: Between the consonant and start of fadeout
Overlap: ⅓ to half of preutterance
Fixed: A good chunk of the fading consonant (approx. 2-3 times the preu.)

Cutoff: Tail end of waveform

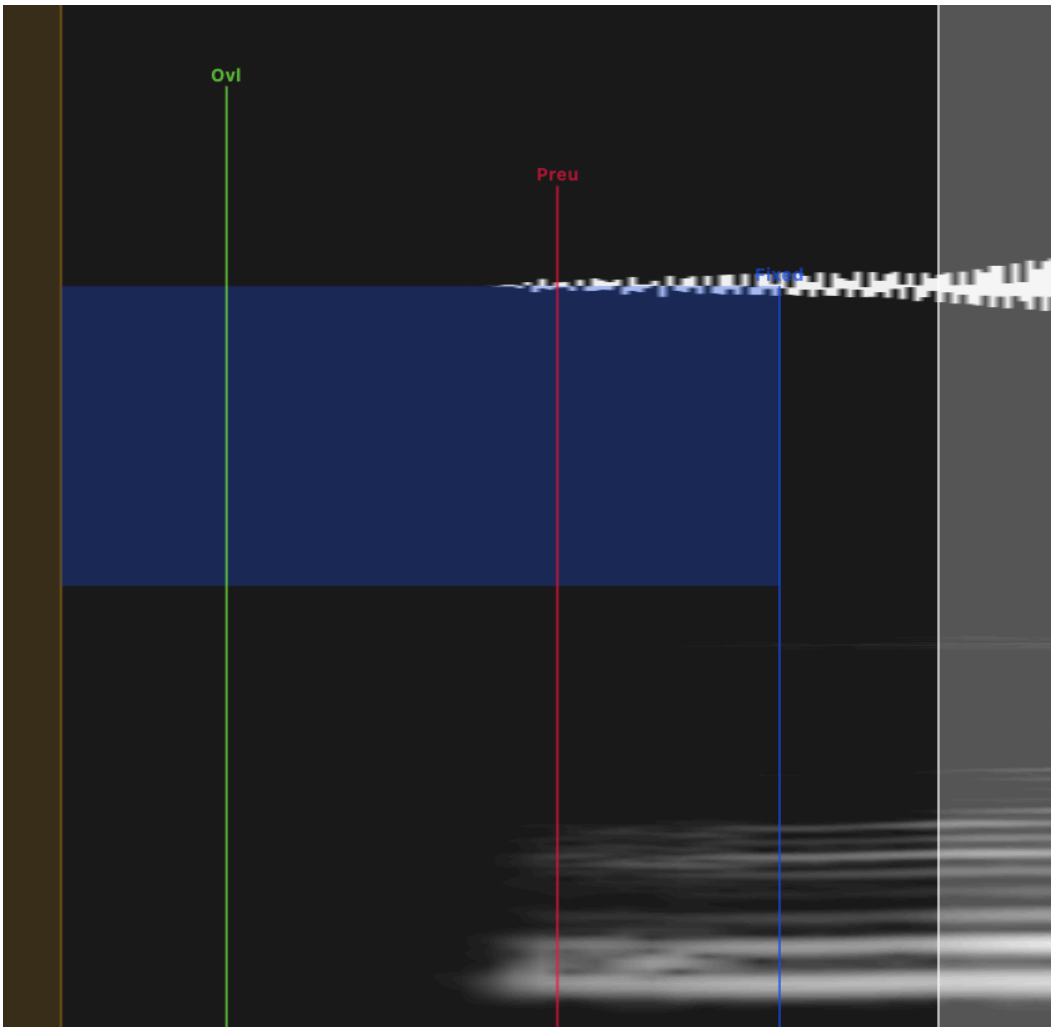



iv. Semivowels (r/rr, w, y)

If you want to use your voicebank with OG UTAU, you would need to use [rr] aliases and exclude [r], since [r] registers as an end breath. For OU, [rr] and [r] should be otoed the same.

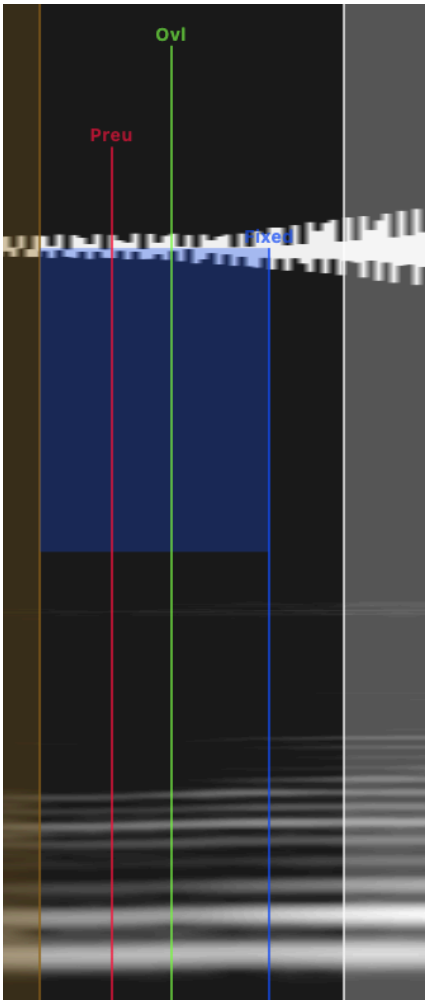
[- C] Starting

**Preutterance:** Start of consonant
Offset: Silence before consonant
Overlap: ⅓ of preutterance
Fixed: When the consonant starts to be stable
Cutoff: Before the consonant starts to blend with the vowel



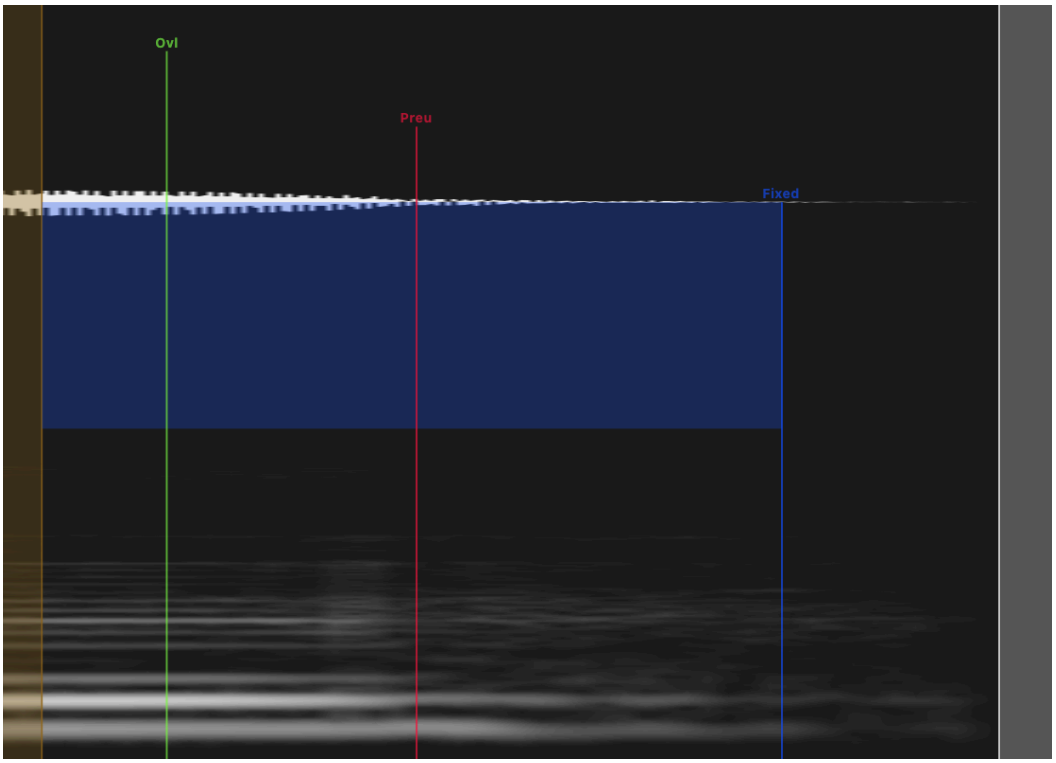

[C] Middle

**Preutterance**: When consonant is stable
Offset: When consonant starts to become stable
Overlap: Twice the preutterance
Fixed: Until the consonant is stable in pitch/volume
Cutoff: Before vowel starts




[C -] Ending

**Preutterance**: Between the consonant and start of fading consonant
Offset: When second consonant is still stable
Overlap: ⅓ to half of preutterance
Fixed: A good chunk of the fading consonant (approx. 2-3 times the preu.)
Cutoff: Tail end of waveform




## c. Vowel Blends and Extras [V C] and [V1]

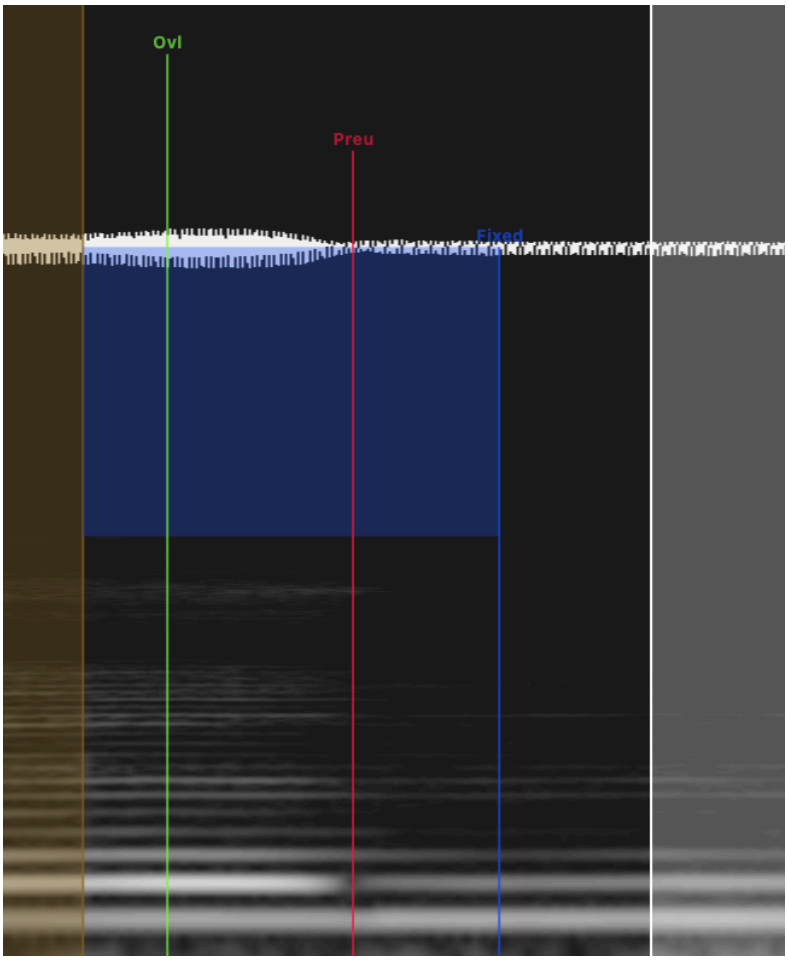
There are some additional oto entries in the base oto that help make synthesis smoother. These Vowel blends will be used by the phonemizer.

You can also oto the vowels present in these samples as alternate takes, like Veria's [V1]s and [V2]s (EX: ae, ae1, and ae2). These alternate takes will only be starting and/or middle aliases. This is optional if you want the most basic c+v eng experience, and will not be used by the phonemizer.

i. Regular Vowels - Consonant (aa_r, ae_n, ae_ng, ao_r, eh_r, ih_ng, ih_r, uh_r, uw_r)

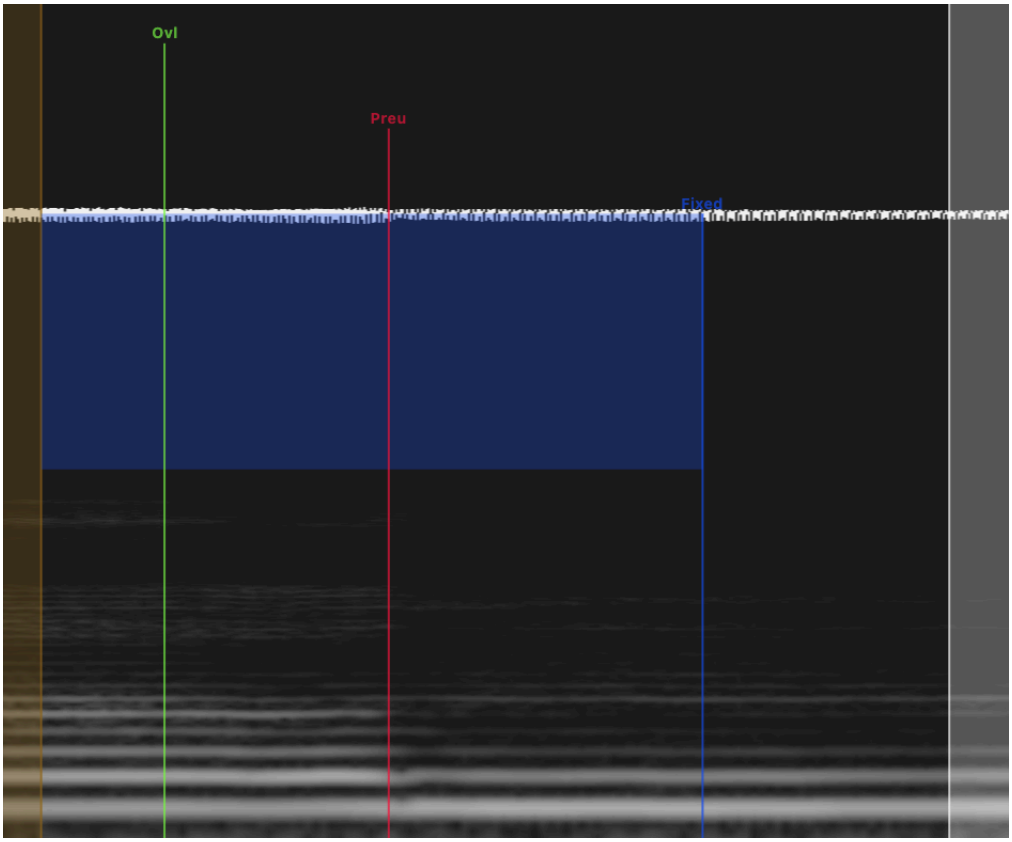
**Preutterance:** Between end of vowel and start of consonant
Offset: When vowel is still stable
Overlap: ⅓ of the preutterance
Fixed: When the consonant gets stable in pitch/volume

Offset: Twice of fixed/until consonant is stable




ii. Diphthongs - Consonant (aw_n)

**Preutterance**: At the end of the second vowel (w)
Offset: When first vowel is still stable
Overlap: At the start of the second vowel
Fixed: When the consonant gets stable in pitch/volume
Offset: Twice of fixed/until consonant is stable




# 3. USING C+V ENG IN OPENUTAU

## Setting Up Voicebank

Your voicebank is now ready to use! I would suggest finishing setting up your voicebank at this stage. Here are some things that OU can do that the guide (that follows OG UTAU) can't:

1. Supporting character icons bigger than 100x100 px!
2. Supporting character images in the piano roll!
3. Setting a default phonemizer from the get-go!

All of these can be done by going to Tools -> Singers -> Cog icon⚙️

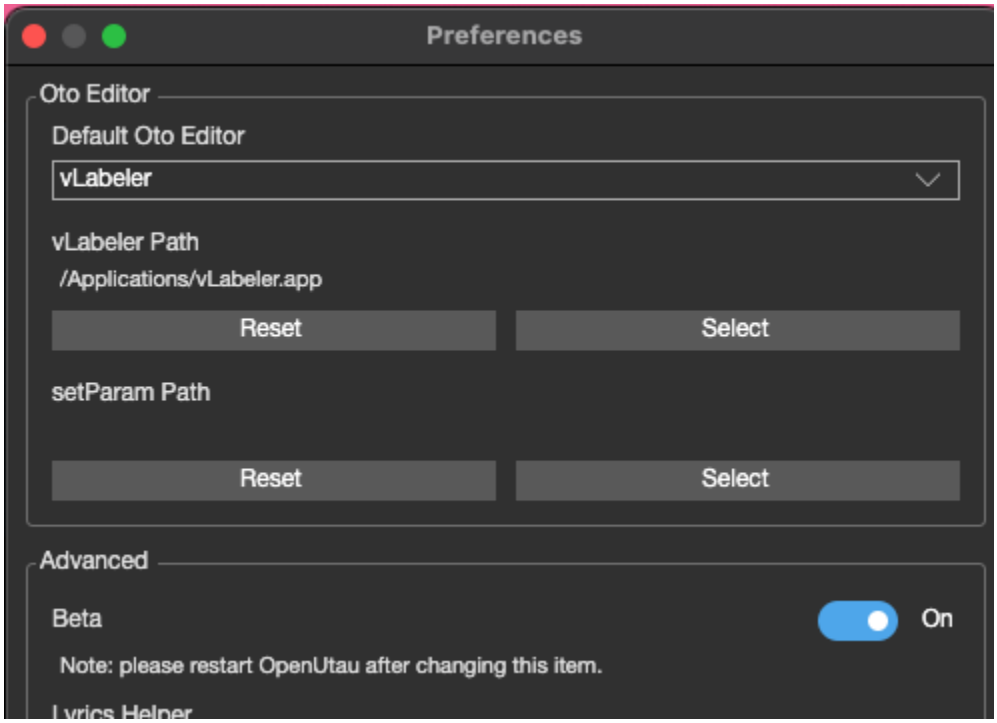
You also might want to generate .frq files.
Make sure to clean up all the .lbp/cache files before zipping and sharing to other people.

# Cadlaxa's C+V ENG Phonemizer

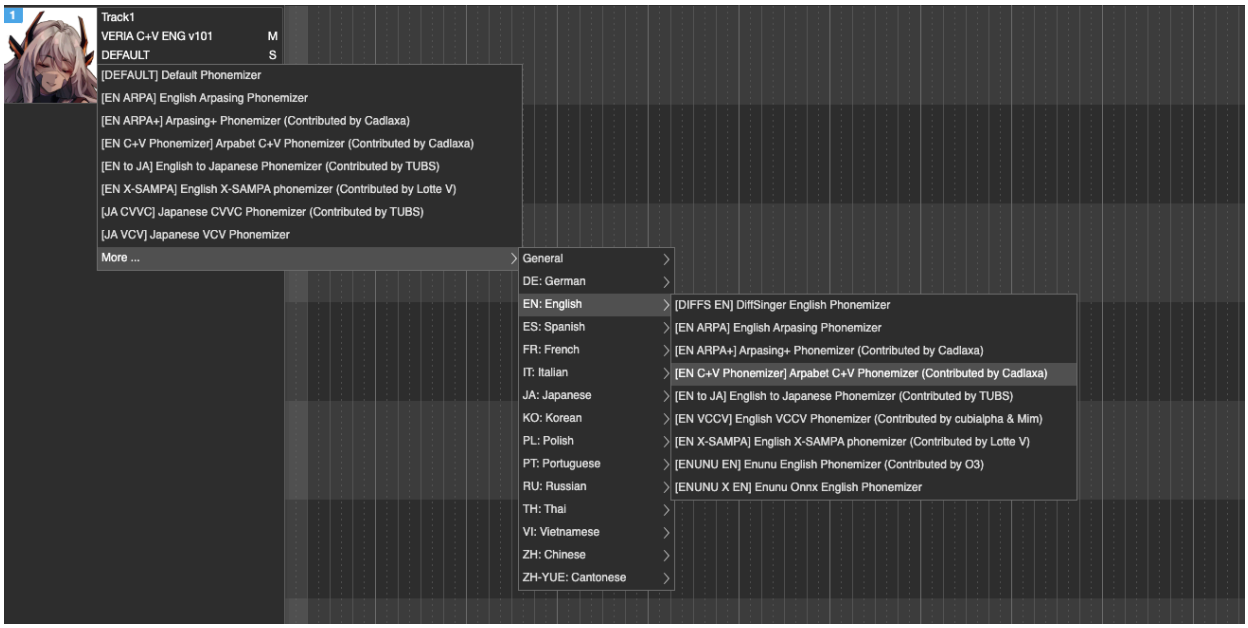
Thanks to Cadlaxa, C+V English now has a dedicated phonemizer! This phonemizer was created with Biggity and another C+V English format (Astel's) in mind, and will be fully compatible with your vb if you follow this tutorial.

The phonemizer is now integrated into OpenUTAU version 0.1.549 Beta! If you have downloaded the external phonemizer before and want to use the integrated version (which has improvements and multilingual support), please delete the external phonemizer BEFORE updating.

To access the Beta version of OpenUTAU, go to Tools -> Preferences -> Advanced, and switch Beta to On.




After installing and restarting OU, you can find the phonemizer underneath the singer's name, in the English section.




## EN ARPA -> Dedicated C+V

If you followed the previous version of the guide that was made prior to the creation of the dedicated phonemizer, you will need to revert back your voicebank to its original state.

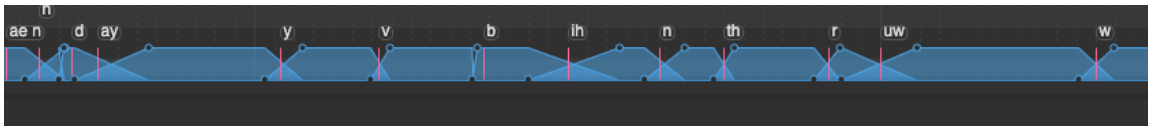
1. Open the location of the voicebank and go into the sample folder. Drag the second folder (ARPA_PLUS) outside of the folder (BASE), and delete the original folder. Rename the remaining folder accordingly (If unsure, use BASE).
2. Open Vlabeler through OU and check that you are in the oto with the suffixes (look at _1). Go to tools -> batch edit -> batch edit entry prefix/suffix. Suffix, remove, _1.
3. Save. Go back to OU and in singers -> edit subbank, and remove the ARPAPLUS colour.

## C+V Eng Usage And Tuning

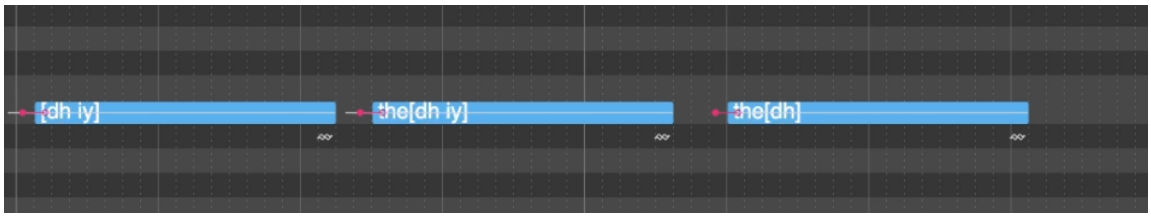
You now have your C+V ENG voicebank ready to use. Here are some pretty important tips when working with a C+V ENG bank.

(Note: Some of these screenshots are from the previous version of the guide, so the phonemes may not match up with how the current phonemes look with the phonemizer!)

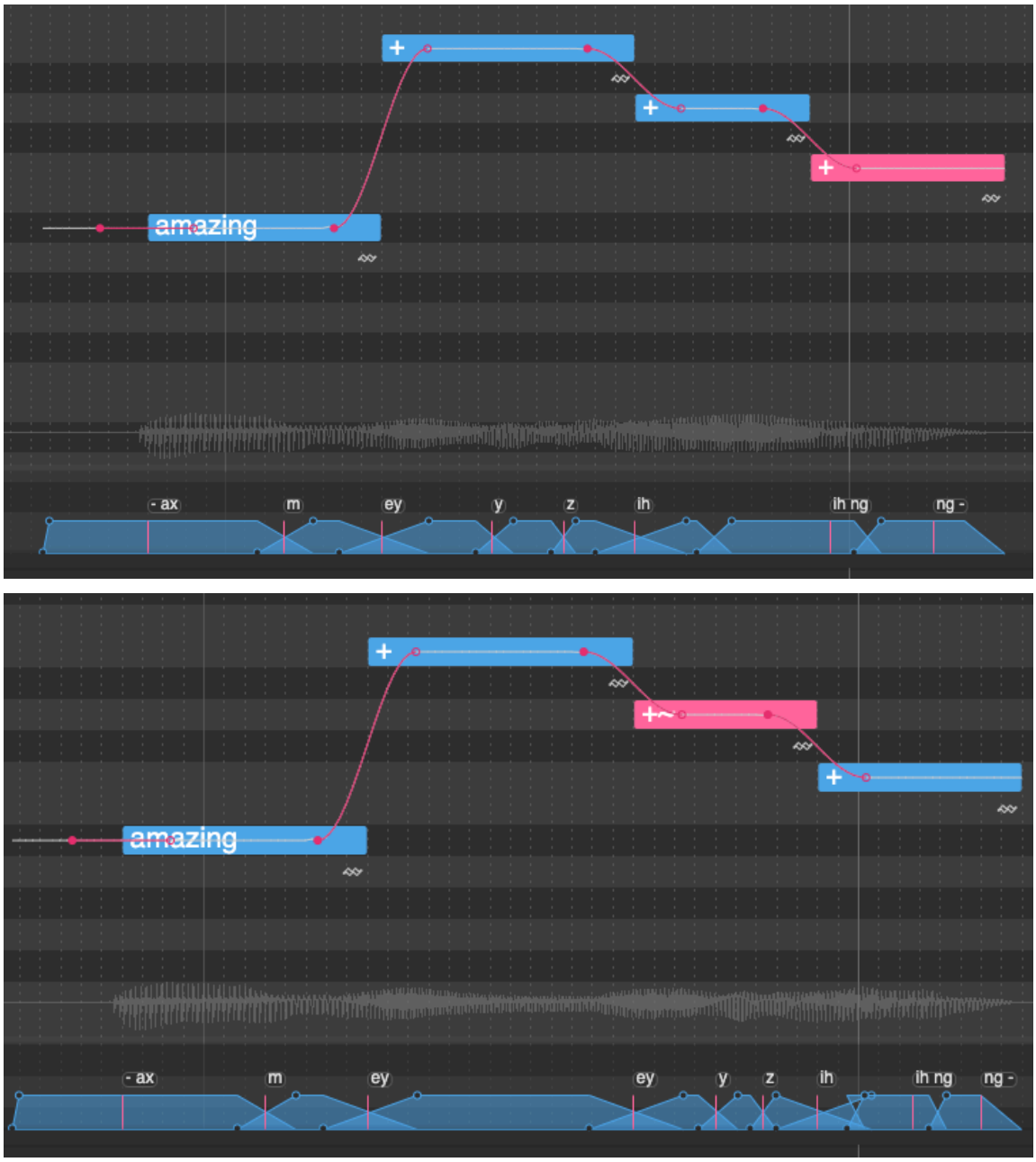
1. You will need to play around with your words' phonemes a lot. Familiarize yourself with the Phonemes tab (toggled with O on the piano roll), where you can edit the crossfade between phonemes, slide where the phonemes start, and edit the phonemes used.



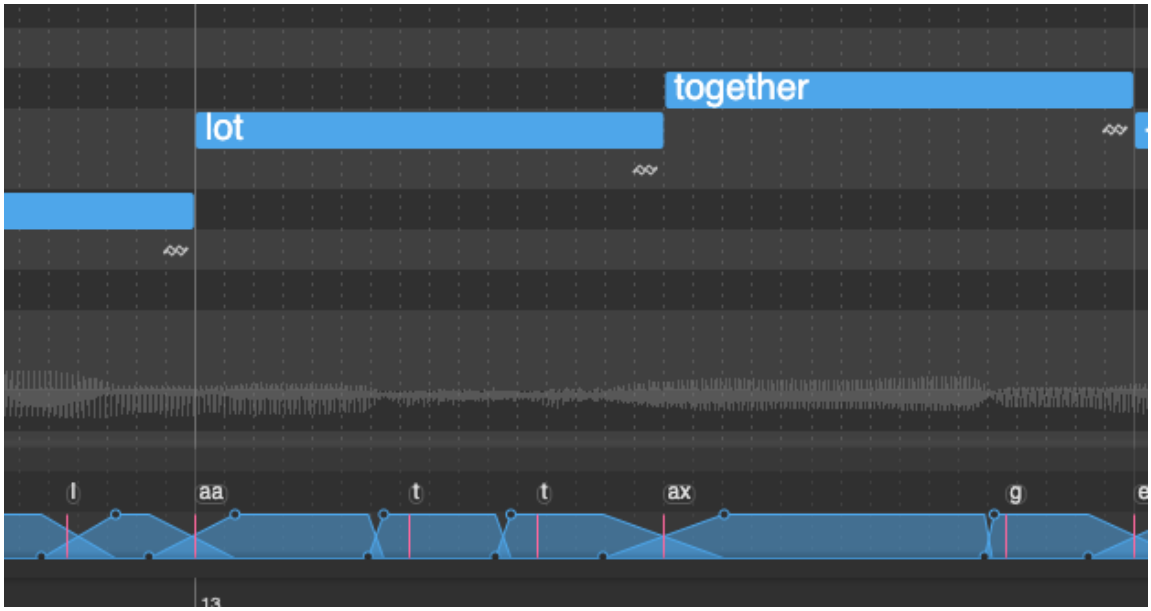

2. On the topic of editing phonemes, you can change your phonemes on the piano roll using [ ].
   Ex: "The" is phonemized as '[dh] [ax]'. You can change it to '[dh] [iy]' by writing The[dh iy] in the lyric OR just [dh iy]. If you want to delete the vowel and just put it as '[dh]', this is also how you do it.

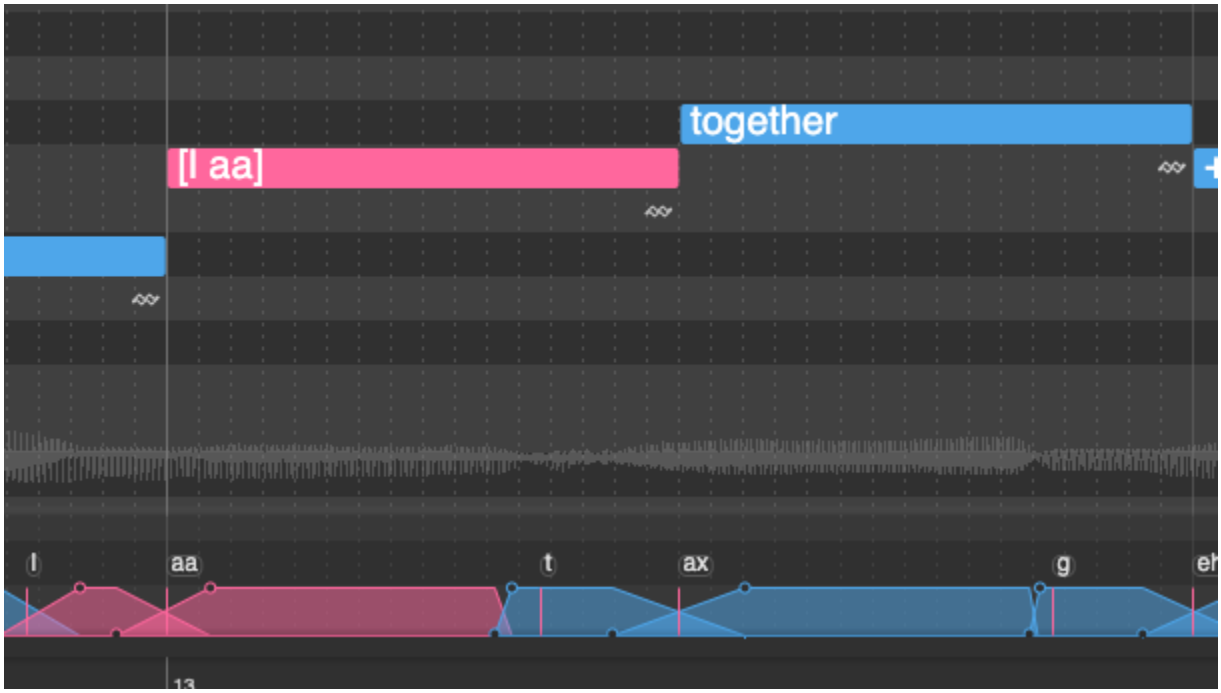



3. To break up a word in multiple notes, use +. If you want to hold a section of the word, use +~. +* also works.




4. The phonemizer will sometimes synthesize two hard consonants one after the other and create a stuttering sound. Edit one of the notes to get rid of the double hard consonant.




# 4. C+V PHONEMIZER FORMATS AND MECHANICS

This entry was contributed entirely by Cadlaxa.

These are the phonemizer mechanics and all the currently supported phoneme formats.
**Note from the phonemizer developer:** The current phonemizer mechanics can be seen within this pull request and changes to the behavior will make the external phonemizer to be outdated, see the link here [Add [EN C+V Phonemizer] by Cadlaxa · Pull Request #1293 · stakira/OpenUtau (github.com)]

## a. Phoneme Fallbacks

The phonemizer tries different alias formats on the voicebank and uses the available format that matches on the phonemizer code. It will try and try until it finds an available phoneme, else, it will not add the phoneme if all of the current possible formats are not met.

## b. Starting Vowel

This is where the starting vowel of the voicebank will be seen on the alias box, the phonemizer tries [-V] -> [- V] -> [_V] then [V]

## c. Vowel-Vowel Transition

The basePhoneme or default phoneme for VV is [prevV V]. But if there's no [V V] aliases found and the transition is a diphthong, It will split them apart ([V C] + [C V]).

1. **Splitting Mechanics (VC part)**
   basePhoneme = $"{prevV} {vvExceptions[prevV]}" (eg: [ay y], [ow w], [ey y]). If the first basePhoneme vc is not available, it will try the other following formats: [V-] -> [{vvExceptions[prevV]}](the w's and y's) -> [V -] -> [_V] -> [-{vvExceptions[prevV]}] then [- {vvExceptions[prevV]}]
2. Splitting Mechanics (CV part)
   Uses only the V's because the VC's takes the consonant part of the CV
   [-V] -> [V] -> [_V] then [- V]

## d. Starting CV With One Consonant (CV)

- Tries all of these vowel sequence [-V] -> [V] -> [- V] then [_V]

- It will also tries to add the consonant part which are [- C] -> [-C] then [C]

## e. Starting CV With More Than One Consonant (CCV/CCCV/C++V)

- Tries all of these vowel sequence [-V] -> [V] -> [- V] then [_V]
- If there are CC's available on the voicebank, it will try to use them first before using singular C's
- CC's: {string.Join("", cc.Take(i))} = means it will concatenate the cc's into one alias like [d r] to [dr] [s k w] to [skw], tries [-CC] -> [-CC] then [CC]
- C's: if the voicebank don't have CC's on them, it will only use singular C's following by a C consonant clusters [- C] -> [-C] then [C]

## f. Middle CV and CCV

Same mechanics with the starting CV's and CCV's

## g. Vowel-Consonant Transition

- First, if the transition has a diphthong, it tries [V-] then [_V] as a VC then adds starting C's to them, [C] -> [-C] then [- C]
- If it's still a diphthong transition and [V-]/[_V] is not available, it will use [y/w] instead with the same starting C's fallback
- If the voicebanks has [V C] ({prevV} {cc[0]}), it will use them, else only the starting C's

## h. Consonant-Consonant Transition

Uses [C] -> [-C] then [- C], also [CC] -> [-CC] then [- CC]

## i. Vowel Endings

Tries [V -] [V R] then [V-]

## j. Consonant Endings

- Same mechanics with the VC + [VCC] and [VC C] support
- Certain consonants are doubled for timing purposes such as the affricates, nasals, liquids, and semi-vowels

# 5. ADVANCED TIPS

## a. Multipitch

You can record multiple pitches and assign them to your voicebank using the Edit Subbanks menu (Tools->Singers). If you are doing multipitch, I would recommend using a GuideBGM.

Some phonemes, like hard and soft unvoiced consonants can be skipped when doing multipitch voicebanks, as they don't carry a pitch with them. You can just reuse the same recordings and oto on all its pitches.

You can reference how a multipitch C+V English voicebank is formatted through Ceta Nocturne.

If you are interested in recording a multipitch voicebank, there is a dedicated toolkit available for you that contains:

1. Separated reclist for unpitched and pitched phonemes
2. Splitted otos for unpitched consonants and other phonemes
3. Template pitch folder setup

After recording and otoing all the files, make sure to go to Tools -> Singers -> Edit subbank and assign the pitches accordingly.

Sometimes OU will fail to detect multiple pitch folders. If so, putting a blank oto file in the main folder should help.

## b. Multilanguage

Due to how C+V is recorded and configured, you can in theory combine voicebanks to create a multilanguage voicebank. This adds more sounds that aren't covered by the original recordings.

Ex: Veria has a C+V English and CV Japanese voicebank. C+V ENG doesn't have the japanese 'R' sound. By combining the two

voicebanks together and creating a combined oto file, Veria can use the ‘Ra, Ri, Ru, Re, Ro’ phonemes if needed.

# 6. FAQ

Q: I have a question/I want to troubleshoot __.
A: You can send your question to my socials or to my contact form! Please read the other entries below for common questions before asking.

Q: This tutorial oversimplifies __!
A: This tutorial is created for total beginners in mind - beginners who just started using UTAU or have only made a CV JPN voicebank. Certain topics and phrases are simplified to help these people understand UTAU terminology.

Q: Did you base Veria/this tutorial on (voicebank)?
A: Absolutely not. My team did not base our method on any other existing voicebanks with the same format(CV/C+V). We would never disrespect previous CV ENG creators' wishes to not reverse engineer their vb.

Q: Would this tutorial work for original UTAU voicebanks?
A: It does! You will miss out on the phonemizer if you use OG UTAU though. If you are going to use it with OG UTAU, make sure your oto file is converted into SHIFT-JIS, because all of Veria’s files are UTF-8.

Q: Should I use a GuideBGM?
A: If you are doing a multipitch voicebank, yes! Otherwise you are free to not do so. Veria was not recorded with a GuideBGM, but she is made with exclusively monopitch in mind.

Q: Can I use another reclist/Can I modify Veria's reclist?
A: Sure you can, but I don’t guarantee it will be fully compatible with the phonemizer! If you are basing the reclist from Biggity’s, please write “based on Biggity_boy’s C+V ENG reclist” to credit his work.

Q: I don't use VLabeler, can I do this with SetParam?
A: Yes, but the terms used will be different. I have never used SetParam, so I can’t say much about it.

Q: Do you have a guide for X-SAMPA/VCCV/etc. formats?
A: No. I am only familiar with ARPAbet. Feel free to make one if you are interested though!

Q: MEL/SEL/etc. is easier in my opinion!

A: That's great! I acknowledge that there are other formats for English synthesis that are known to be simple that are more established than C+V. This format is simplest to me, the creator of this guide - someone who was intimidated by non-ARPABET phonemes and very long vb production times. The claim 'English UTAU for Total Beginners' is not meant to and should not be taken as an offense to other formats and their enjoyers. It's all in good fun:o)

Q: I saw the previous video/version of this doc, and the oto guide is completely different. Why's that?
A: The previous contributor has requested to pull all their work from this guide. In addition to that, I also incorporated user feedback and phonemizer usage into the guide, and this current version of the oto guide will net you the best results with the dedicated phonemizer.

# 7. PHONEME LIST

Imported from Veria's now retired user guide.

**Standard American English pronunciation. Based on the SynthV phoneme chart.**

| VOWELS | | | |
|---|---|---|---|
| **Phoneme** | **Example** | **Phoneme Classification** | **Vowel Blends** |
| aa | p**a**lm | Regular | aa r |
| ae | b**a**ng | Regular | ae n, ae ng |
| ah | b**u**t | Regular | |
| ao | b**ough**t | Regular | ao r |
| aw | ab**ou**t | Diphthong | aw n |
| ax | rent**a**l | Regular | |
| ay | b**i**te | Diphthong | |
| eh | b**e**t | Regular | eh r |
| er | b**i**rd | Regular | |
| ey | h**ey** | Diphthong | |
| ih | b**i**t | Regular | ih ng, ih r |
| iy | b**ea**t | Regular | |

| | | | |
|---|---|---|---|
| ow | boat | Diphthong | |
| oy | boy | Diphthong | |
| uh | book | Regular | uh r |
| uw | boot | Regular | uw r |

| CONSONANTS | | | |
|---|---|---|---|
| Phoneme | Example | Phoneme Classification | Start and End Aliases and Comments |
| b | buy | Hard | |
| ch | china | Hard | |
| d | die | Hard | |
| dx | butter | Soft Voiced | Only start and also ‘ suffix on the notes that ends with t/d turns them to dx |
| dh | the | Soft Unvoiced | (Struggled the most with recording this. May be the least accurate sounding phoneme) |
| f | fight | Soft Unvoiced | |
| g | guy | Hard | |
| hh | high | Soft Unvoiced | Can be used as ending exhales |
| jh | just | Hard | |
| k | kite | Hard | |
| l | lie | Soft Voiced | |
| m | my | Soft Voiced | |
| n | nose | Soft Voiced | |
| ng | sing | Soft Voiced | |
| p | pie | Hard | |

| q | (Glottal stop) | Hard/Special | Phonemizer will use q with the use of ‘ prefix |
|---|---|---|---|
| r/rr | right | Semivowel | |
| s | say | Soft Unvoiced | |
| sh | shy | Soft Unvoiced | |
| t | tie | Hard | |
| th | thing | Soft Unvoiced | |
| v | vote | Soft Unvoiced | |
| w | wise | Semivowel | Only start |
| y | yes | Semivowel | Only start |
| z | zoo | Soft Unvoiced | |
| zh | measure | Soft Unvoiced | |

| EXTRAS | | |
|---|---|---|
| Phoneme | Example | Note |
| br | breath | Very sharp, short breath |

| MISSING CONSONANTS AND SUBSTITUTIONS | | |
|---|---|---|
| Phoneme | Example | Substitution |
| dr | drive | [d] [rr] or [dh] [rr] (phonemizer fallback: [jh] [r]) |
| tr | tree | [t] [rr] (phonemizer fallback: [ch] [r]) |

# 8. CREDITS

a. C+V ENG Project
   i. Project Lead, Guide Writer, Video Editor - chevrefee (twitter) (wordpress) (youtube)

ii. Contributor, Reclist creator, Base oto provider - Biggity_boy (twitter) (youtube)
iii. Contributor, Oto revision input - Anemonie (youtube) (soundcloud)
iv. Contributor - Vyreed (twitter)
v. Oto revision input - Cytalai (twitter) (website)
vi. Phonemizer creator, guide editor, oto revisions - Cadlaxa (twitter)
vii. Sample providers for revised oto beta testing - Vyreed (Teo Vampa), Gravekind (Wakae Sota)

b. Software
   i. OpenUTAU - STAkira
   ii. VLabeler and RecStar - sdercolin
   iii. OREMO - nwp8861
c. Recording Section
   i. Phoneme guide - Dreamtonics
   ii. Alternate recording - Biggity_boy
d. Otoing Section
   i. Otoing terms, Arpasing, CVVC JPN, VCCV ENG oto technique - Yin-P
   ii. Original Oto technique - Biggity_boy
   iii. Adapted Oto technique - KLAD via Arpasing neocities website
   iv. Adapted Oto technique - Cabbagesaurus
   v. Adapted Oto technique - Studio VOXYZ
   vi. Adapted Oto technique - Cytalai
   vii. Adapted [V-] diphthong oto technique - Astel
e. Usage Section
   i. Voicebank setup - utaututorials
f. Phonemizer Format Section
   i. Entire section - Cadlaxa
g. Advanced Tips
   i. Multipitch setup - Vyreed
h. Video
   i. Spanish closed captions - Gravekind
   ii. German closed captions - Hyazinthenfuchs
   iii. Ukrainian closed captions - phi_pea and the Vocalsynth UA team
   iv. Swedish closed captions - WasThatZero
   v. Indonesian closed captions - emptyshelled

# 9. CHANGELOG

Apr 13, 2025
Deleted external phonemizer section, replaced with integrated phonemizer usage, added Indonesian closed captions credits

Oct 24, 2024
Added reference for multipitch voicebank

Oct 16, 2024
Added credits for Swedish closed captions

Oct 14, 2024
Edited diphthong oto section [V] and [V-], added disclaimer about base recording differences, changed glottal stop entry in phoneme list

Oct 12-13, 2024
Edited links for external phonemizer, added credits for Ukrainian closed captions

Oct 9, 2024
Edited the phoneme list entry for q, modified diphthong oto section and added [V-] subsection, added credits for the oto

Oct 8, 2024
Added credits for German closed captions

Oct 7, 2024
Added C+V phonemizer formats section

Oct 5, 2024
Removed previous contributor’s work, overhauled guide - added more details for recording, complete overhaul of oto guide, removal of EN ARPA configuration, merged usage tips and OU usage to one section, added more details on multipitch subsection, added full credits section, merged Veria’s user guide with this document and importing the phoneme list, added more contributors, properly credit oto techniques, updated links

Sep 22, 2024
Published document

Sep 17, 2024
Created document