EASYTALK™

DOCUMENTATION

CONTENTS

1. OVERVIEW

EasyTalk brings your Unity game characters to life with lip syncing, facial expressions, head, and eye contact. It was designed from the ground up to be easy to use and intuitive. With EasyTalk, you can create animations quickly. Once you have some audio clips and a character set up in Unity, you can literally put together fully lip synced, facially animated sequences in just minutes.

2. GETTING STARTED

Use the Unity Asset Store download manager to download and import the EasyTalk asset into your project. You should not move this asset to another folder. It will be located in your Project's /Assets/EasyTalk folder.

Before you dive into the documentation, it is recommend that you watch the quick start tutorial video available on the website at http://www.realtimevr.com.

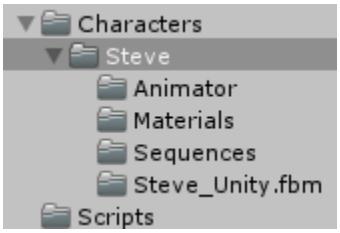
There is a sample demo scene in the /Assets/EasyTalk/Scenes/DemoScene.

3. THE WORKFLOW

EasyTalk requires any 3D model with viseme, and optionally expression, blendshapes. If you create a character in DAZ Studio or Autodesk Character Generator, these blendshapes are automatically created for you. You are not restricted to these authoring tools but EasyTalk supports them out of the box.

3.1. PREPARE THE UNITY PROJECT

It is recommended that you create separate folders for each character asset in your Unity project along with a "Sequences" folder attached to each character. This will help keep things organized and give you a place to export sequences to later.

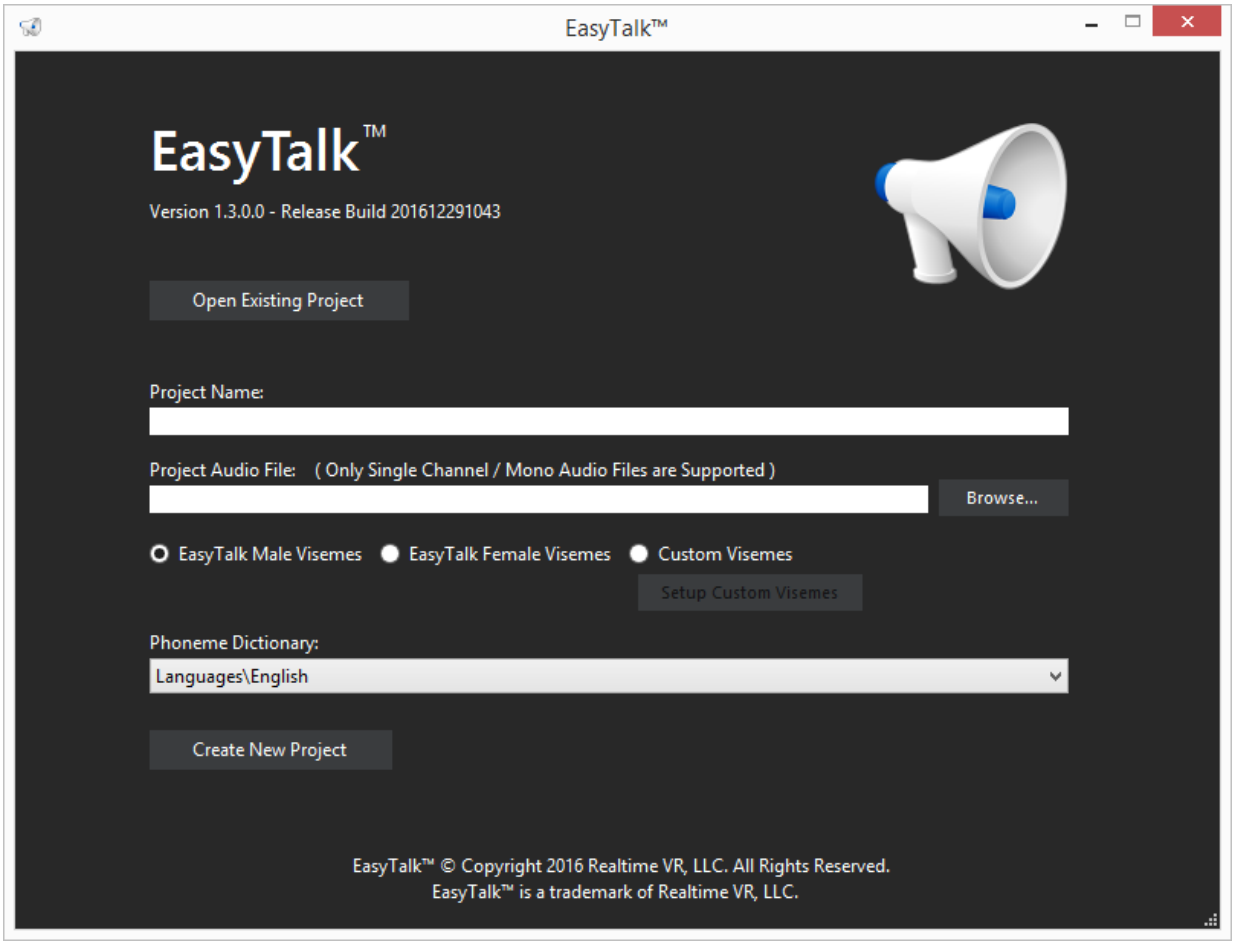


3.2. CREATE A LIP SYNC'D SEQUENCE

If you have not done so already, you need to download and install the EasyTalk Windows application. The application can be downloaded from our website at http://www.realtimevr.com or directly from http://www.realtimevr.com/EasyTalkSetup.exe - This application is 100% free to download.

3.2.1. CREATE AN EASYTALK PROJECT WITH AN AUDIO CLIP

When you start the EasyTalk application you will be presented with the open/create project window:

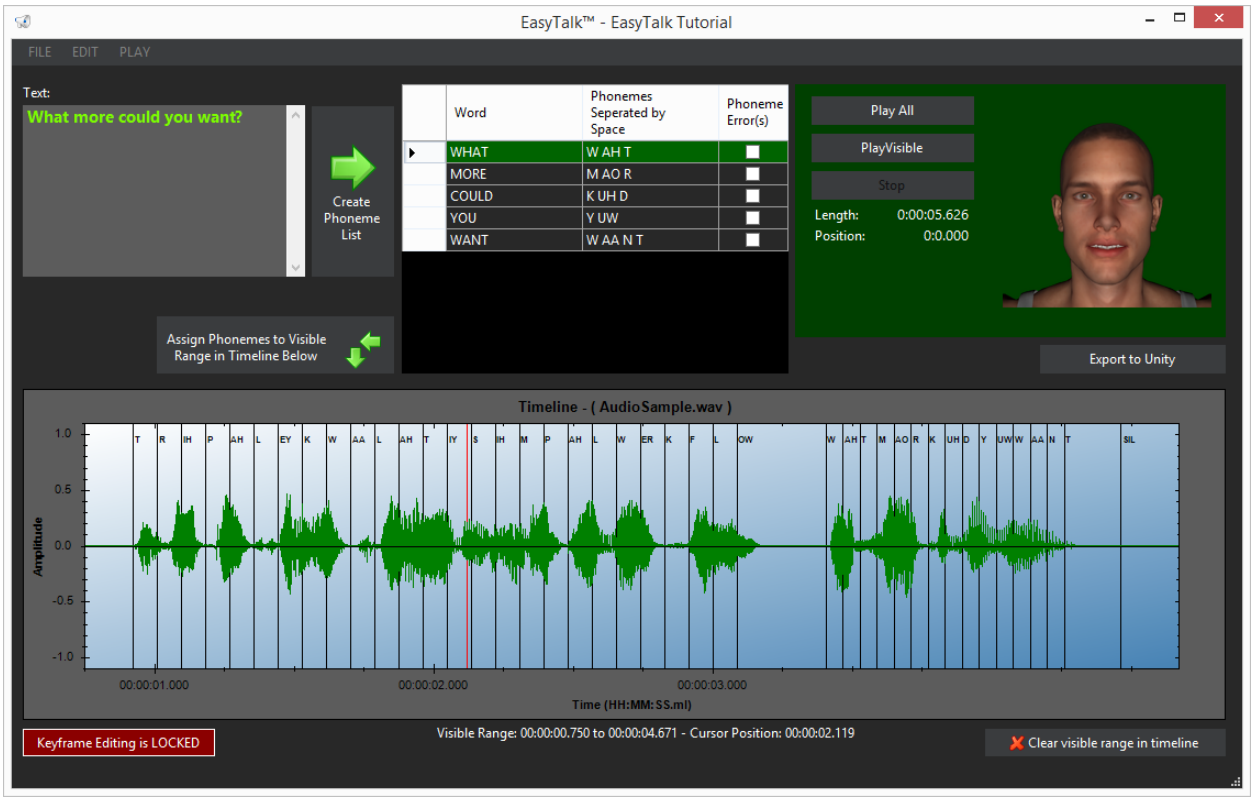


Audio clips of your speech must be mono not stereo. You can convert stereo files to mono using Audacity which can be downloaded for free online.

Audio clips should be have a maximum duration of about 45 seconds. Going beyond that will degrade performance. You can always play back sequences consecutively in your game to address this.

The “EasyTalk Male Visemes” and “EasyTalk Female Visemes” radio buttons initialize the project with default male or female character viseme shape images to help you preview your sequences. If you are using custom visemes, select “Custom Visemes” and choose the viseme images you created for your character. These viseme images are expected to be 190x190 pixels in size but if they are not, they will be scaled automatically.

The English language, and a subset of French, is supported. Select the language of choice and create the project. Once you have created a project, the main application window appears as shown here:

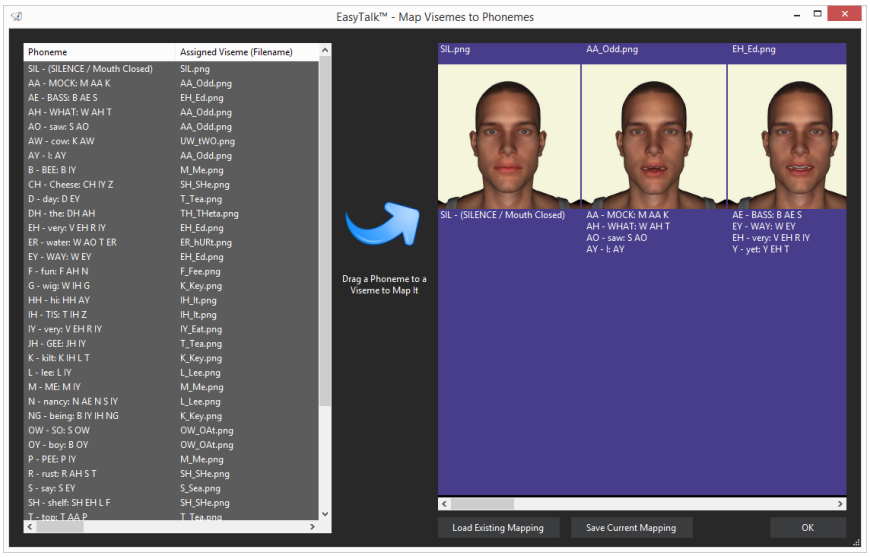


Simply type the words into the “Text” area and click on the [Create Phoneme List] button. The words and their corresponding phonemes will appear in the section to the right. If any words are not in the dictionary, the words will be highlighted in red. You can add or edit dictionary words by clicking on them.

3.2.2. Map Visemes to Phonemes

If you need to map your visemes to phonemes, or if you are using custom visemes, you will need to map them.

The processes of mapping visemes to phonemes is rather straight forward. Select “Map Visemes to Phonemes” from the EDIT menu.



3.2.3. Use the Audio Timeline

The best way to place phonemes into the timeline and the best way to get great results without having to drag phoneme markers around, is to first zoom in on a discrete subsection of audio. You can do that by dragging a rectangle around a section of audio in the timeline. Before you attempt to zoom in on the audio section, ensure that the [Keyframe Editing is Locked] indicator is enabled (red) or you may inadvertently drag a phoneme marker. The [Keyframe Editing is Locked] indicator is also a button. Click on it when needed to enable or disable keyframe editing.

If you need to zoom out of a section in the timeline, simply right-click on the timeline and select “Un-Zoom” from the popup window.

You can playback the entire audio clip or just a visible subsection in the timeline using the playback buttons. Homing in on a few seconds of audio is best. By using the waveform in the timeline as a guide, you can get a good idea of where a group of words start and end.

3.2.4. Place Phonemes Automatically

Phonemes are always placed on the currently visible section of the timeline. Zoom into the timeline section you wish to place phonemes into.

You can manually place phonemes onto the timeline at any time by right-clicking in the timeline. You can delete phonemes by zooming into the section you wish to clear and clicking on the [Clear visible range in timeline] button.

To automatically place phonemes onto the timeline, you can simply type in the words into the “Text” area, click on the [Create Phoneme List] button, and then click on the [Assign Phonemes to Visible Range in Timeline Below] button. Phonemes are then placed into the timeline.

Again, working on small subsections of the timeline is best and you can always tweak the results by dragging the placed phoneme markers to the desired location. You just need to be sure that Keyframe Editing is enabled with the toggle button.

Inserting an SIL (Silence) phoneme in the middle of areas where no speech is heard is usually a good idea. Also, placing an SIL phoneme at the end of the sequence is recommended which will put your characters mouth back into a known state so that new sequences won't abruptly change mouth positions.

3.2.5. Preview the Lip Syncing

You can get a rough approximation, or preview, of the final lip synced results by playing back the audio and viewing the preview window. Although this is obviously not a smoothly interpolated preview, it will give you a very good idea of any timing issues between the phonemes and the audio. If you don't like the way a section looks, you can always zoom into it and fine tune it by dragging the phoneme markers.

3.2.6. Export the Sequence to Unity

When you are satisfied with the audio lip sync sequence, you can export it into the Sequences folder of your Character folder in your Unity project. Click on the [Export to Unity] button to do this.

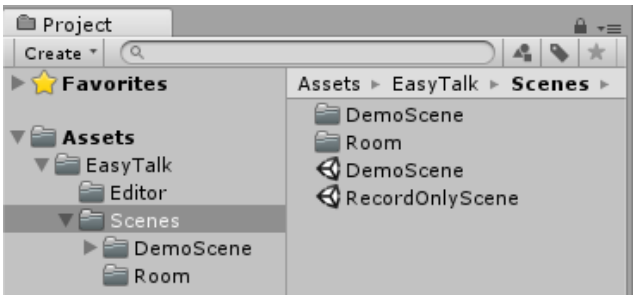
Note that you can export as many Lip Sync Expression sequences as you want for your character into the Sequences folder.

3.3. ADD EXPRESSIONS WITHIN UNITY

With your sequences in hand, you can start working within the Unity editor to add expressions to your sequences.

3.3.1. Set up the Scene

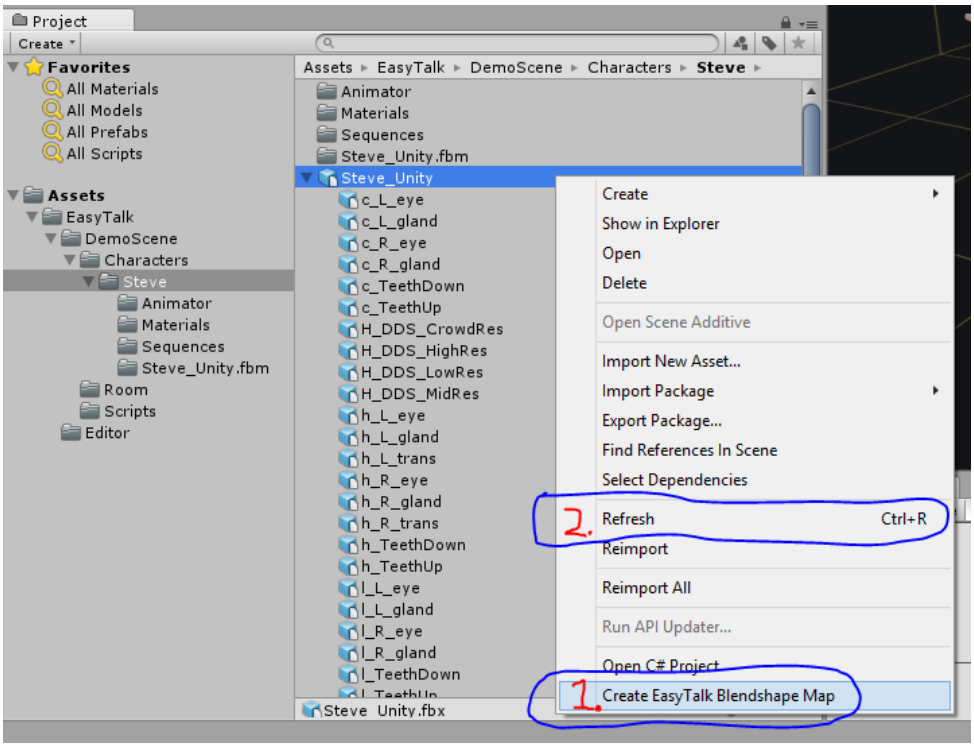
You should create one unique scene in your Unity project dedicated solely to recording expressions and previewing sequences for your characters. The main reason for that is that your input devices, including keyboards, mice, and gamepads, will conflict with the recording mechanism of the EasyTalk script. Fortunately, there is a "Record Only" scene that you can use which is located in /Assets/EasyTalk/Scenes/RecordOnlyScene.



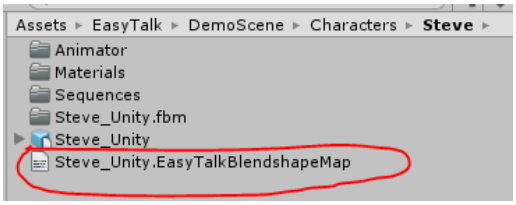
You can drag an instance of each 3D character you have created into this scene for preview and recording purposes.

3.3.2. Create the Blendshape Map (CSV File)

Once you have imported your character model into Unity, right-click on the model in your Assets folder and select “Create EasyTalk Blendshape Map” as shown here:



This will create a Blendshape map in the same directory as the model. Once complete, “Refresh” the window to see the resulting blendshape map file:



This is a CSV file that can be edited either with a simple text editor, or with a spreadsheet application like Microsoft Excel. It is used to provide EasyTalk with a mapping of visemes and expressions to blendshapes.

	A	B	C	D	E	F
56	Expression	OpenMouthSmile	H_DDS_LowRes	60	l_expressions.LLbrowUp_l	46
57	Expression	OpenMouthSmile	H_DDS_CrowdRes	57	c_expressions.RRbrowUp_c	30
58	Expression	OpenMouthSmile	H_DDS_CrowdRes	58	c_expressions.RbrowUp_c	10
59	Expression	OpenMouthSmile	H_DDS_CrowdRes	59	c_expressions.LbrowUp_c	17
60	Expression	OpenMouthSmile	H_DDS_CrowdRes	60	c_expressions.LLbrowUp_c	46
61	Expression	EyesClosed	H_DDS_HighRes	45	h_expressions.ReyeClose_h	100
62	Expression	EyesClosed	H_DDS_HighRes	46	h_expressions.LeyeClose_h	100
63	Expression	EyesClosed	H_DDS_MidRes	45	m_expressions.ReyeClose_m	100
64	Expression	EyesClosed	H_DDS_MidRes	46	m_expressions.LeyeClose_m	100
65	Expression	EyesClosed	H_DDS_LowRes	45	l_expressions.ReyeClose_l	100
66	Expression	EyesClosed	H_DDS_LowRes	46	l_expressions.LeyeClose_l	100
67	Expression	EyesClosed	H_DDS_CrowdRes	45	c_expressions.ReyeClose_c	100
68	Expression	EyesClosed	H_DDS_CrowdRes	46	c_expressions.LeyeClose_c	100
69	Expression	ClosedMouthSmile	H_DDS_CrowdRes	19	c_expressions.RlipCorner_c	100
70	Expression	ClosedMouthSmile	H_DDS_CrowdRes	20	c_expressions.LlipCorner_c	100

Column A is the type which is either Expressions, Viseme, English, or French. Column B is either the name of the expression or phoneme. If it is blank, it is ignored by EasyTalk. Column C is the name of the mesh.

Column D is the index of the Blendshape and Column E is the corresponding name of the mesh. The index in Column E is used by the script and the name in Column E is for reference only. Finally, column F is the maximum strength of the blendshape for the given expression or phoneme. Its range is from 0 to 100.

EXPRESSION MAPPING

Expressions are not mapped automatically. To map an expression in the CSV file, copy one or more blendshapes into a new line in the file and place the name you wish to use for that expression in Column B. For example, here's an example of 2 lines in the file that contain an 'EyesClosed' blendshape for an Autodesk Character Generator character:

	A	B	C	D	E	F
61	Expression	EyesClosed	H_DDS_HighRes	45	h_expressions.ReyeClose_h	100
62	Expression	EyesClosed	H_DDS_HighRes	46	h_expressions.LeyeClose_h	100

Note that there is an example in the tutorial video of how to map expressions. There is also a sample blendshape map file in the DemoScene which is bundled with the EasyTalk Unity Asset located in the \Assets\EasyTalk\DemoScene\Characters\Steve\ Steve_Unity.EasyTalkBlendshapeMap.csv.

PHONEME MAPPING

If you auto-generated a blendshape map file for a DAZ Studio character or an Autodesk Character Generator character, the mapping of the phonemes to viseme blendshapes is performed automatically provided that you exported blendshape information.

However, if you are not using a character from DAZ Studio or Autodesk Character Generator, you need to map the phonemes to your blendshape visemes manually. For the English language there are 39 phonemes. They are:

Phoneme	Example	Translation
AA	odd	AA D
AE	at	AE T
AH	hut	HH AH T
AO	ought	AO T
AW	cow	K AW
AY	hide	HH AY D
B	be	B IY
CH	cheese	CH IY Z
D	dee	D IY
DH	thee	DH IY
EH	Ed	EH D
ER	hurt	HH ER T
EY	ate	EY T
F	fee	F IY
G	green	G R IY N
HH	he	HH IY
IH	it	IH T
IY	eat	IY T
JH	gee	JH IY
K	key	K IY
L	lee	L IY

M	me	M IY
N	knee	N IY
NG	ping	P IH NG
OW	oat	OW T
OY	toy	T OY
P	pee	P IY
R	read	R IY D
S	sea	S IY
SH	she	SH IY
T	tea	T IY
TH	theta	TH EY T AH
UH	hood	HH UH D
UW	two	T UW
V	vee	V IY
W	we	W IY
Y	yield	Y IY L D
Z	zee	Z IY
ZH	seizure	S IY ZH ER

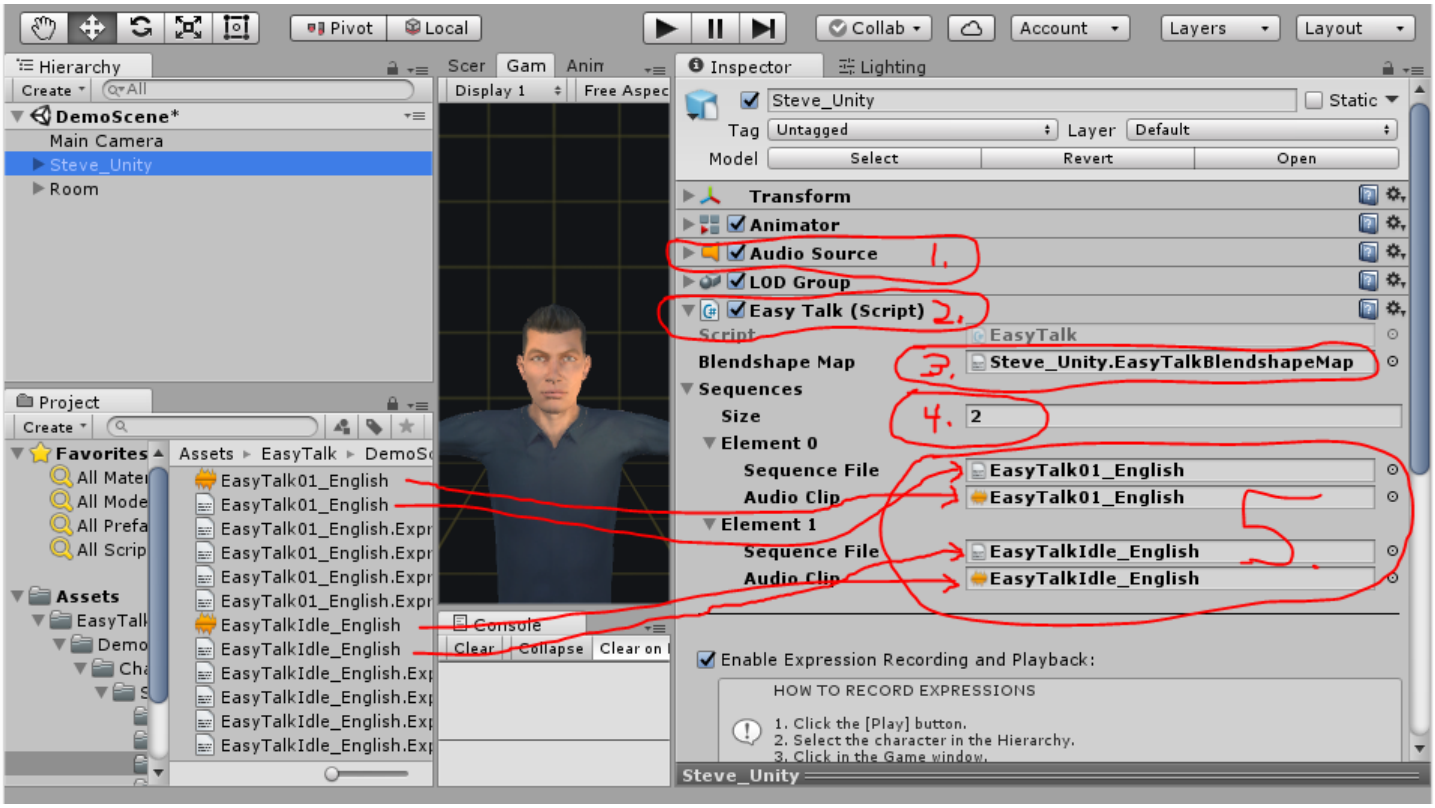
This set is based on the ARPAbet symbol set developed for speech recognition uses.

EasyTalk also adds an SIL “phoneme”. This is to represent a mouth closed (or SILent) “phoneme”. There should be an SIL “phoneme” for each viseme blendshape for your character with a value of zero.

For an example of phoneme mapping, please see the blendshape map file in the demo scene bundled with the EasyTalk asset.

3.3.3. CONFIGURE YOUR CHARACTER INSTANCE

There are basically 5 steps to get things running. Select the character in your scene and perform these steps in the inspector window:

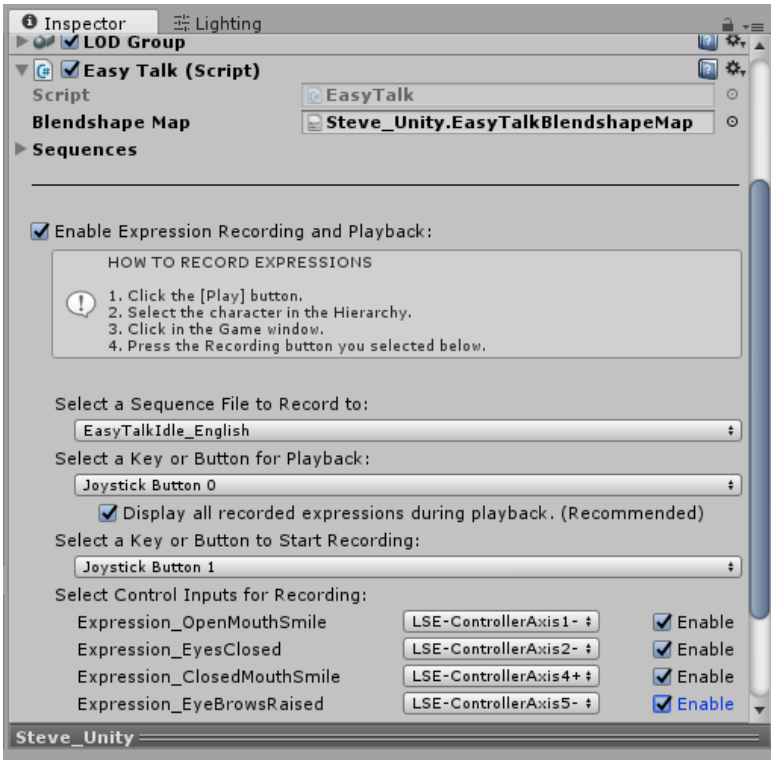


1. Add an “Audio Source” component. You may want to enable 3D audio.
2. Add the EasyTalk script component.
3. Assign the Blendshape Map we created earlier.
4. Enter the number of sequences you wish to assign to the character.
5. Assign the sequence(s) and associated audio file(s). Expand the ‘Sequences’ section and enter the number of sequences you have in the ‘Size’ box. Expand the ‘Element’ section and drag the Sequence File and Audio File that was exported from the Windows app into corresponding boxes.

3.3.4. RECORD EXPRESSIONS

Before you can record expressions, you need to ensure that the expressions are mapped to blendshapes in your Blendshape Map file. For more information, see the section on creating the Blendshape Map file above. If you don’t configure any expressions, no expression will be visible in the EasyTalk inspector panel.

Here’s an example of the expression recording section in the EasyTalk inspector panel:



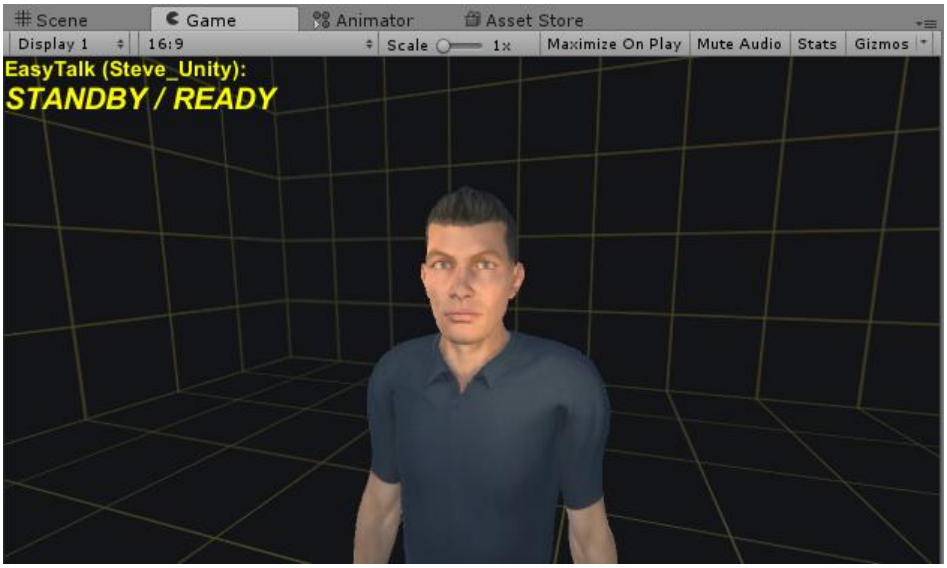
As shown here, there are four (4) mapped expressions.

CONFIGURE SETTINGS AND ENABLE STANBDY / READY MODE

Ensure that 'Enable Expression Recording and Playback' is checked. Select a sequence in the drop down box, and select a button for previewing. Select a different button for recording.

Click on the 'Enable' checkbox next to an expression you wish to record to. Select an input for recording the expression. A joystick with an axis, such as an XBOX controller, is recommended but you can use the mouse to control expressions as well.

Provided that your camera is focused on your character's face in the scene, select your character in the Scene view or Hierarchy first. Now click on the Play button within the Unity editor and you will see your character in the Game window. You should see something like this:



*IMPORTANT: **If there is no yellow text** in the Game window, you need to first select your character in the Scene View or Hierarchy window. If you then see the yellow text but don't see 'STANDBY / READY' indicated in the Game window, you need to click in the Game window. Once you do that 'Standby / Ready' mode is enabled. You should now see text similar to what is shown in the image above.*

PREVIEW THE LIP SYNCING AND EXPRESSIONS WITH THE AUDIO

Congrats! You are now ready to preview your lip synced character. When you are in 'STANDBY / READY' mode, play back your sequence with the button or key you assigned earlier. To preview expressions, move the joystick or mouse axis that you is assigned to it.

RECORD THE EXPRESSIONS

To record an expression, ensure that you are in 'Standby / Ready' mode. Press the record button. Your character will start talking and you can record expressions along with that in real-time. Press the record button again to stop recording.

If you want to do another take on the expression recording, simply hit the record button and start over.

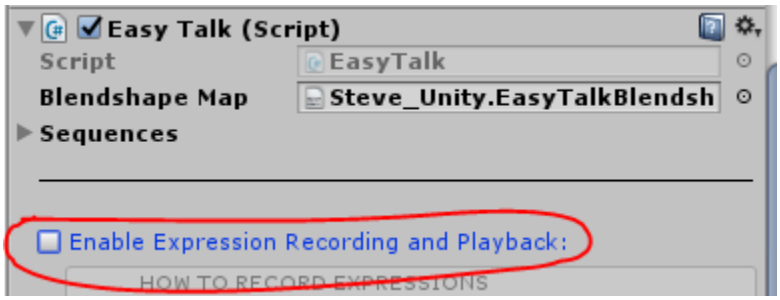
Once you record an expression, it is locked into your animation. Suppose, however, that you decide you no longer want that expression. To remove the expression from the recording, enable the expression in the Inspector, hit record, and simply don't move the corresponding axis.

Recording a particular expression does not affect other recorded expressions. However, you need to disable the other expressions first or they will be overwritten.

*IMPORTANT: Recording a particular expression does not affect other recorded expressions. However, if you want to keep an expression that you've already recorded, **you must disable it** in the Inspector **by unchecking the associated [Enable] check box or it will be overwritten!***

4. PLAYING SEQUENCES IN-GAME FROM YOUR SCRIPTS

To play the sequences in-game, you need to ensure that the EasyTalk script is attached to your character for your scene and that 'Enable Expression Recording and Playback' is disabled.



The following public methods are available to your scripts for playing, stopping, and checking the status of sequences.

- Plays a sequence one time.
  public void PlaySequence(string sequenceName)
- Plays a sequence with optional looping. Works well for idle sequences.
  public void PlaySequence(string sequenceName, bool loop)
- Stops the currently playing sequence.
  public void StopSequence()

- Determines if a sequence is currently playing.
  public bool IsSequencePlaying()

5. SOME HELPFUL HINTS

Experiment and have fun. There's no reason you can't get AAA quality and you are not limited in any way. Once you have setup a character and get the hang of things, you can create sequences in minutes instead of hours. Here's some tips:

- Audio clips should start with a pre-determined amount of silence. This will allow your sequence to "warm up" when playing it in-game. Your character could start with a facial expression, or look at you before talking, etc.
- You may also want to consider using an audio clip that has sighing, coughing, or laughing noises, for example, to add more realism. You can record expressions for those noises.
- The most convincing visemes and expressions are those with imperfections in them. For example, some people raise their right upper lip when they talk. You could curl up that part of the mouth in your visemes or create an expression for that and overlay it on top of your lip synced sequences.
- When people talk, they generally don't have an expressionless face. If someone is happy when they talk, they might have a smile. You should try to add some form of subtle facial expression when they are talking.
- If you are planning on using a particular expression while someone is talking, be sure to keep their mouth closed when you create the blendshape for that expression. Expressions are additive. To expand on that, if you have a character with an open mouth smile, and you combine that with lip syncing, the mouth will be more open than it should be.
- Subtle nuances and imperfections can pay off big in your final sequence.
- Use an eye brow raise expression for when the character talks. This is very common.
- Create mood expressions. If you want to create mood expressions for when a character is talking, simply remove the mouth blendshapes so that they don't interfere with the mouth movements.
- Always add blinking.
- Create an idle sequence with at least some blinking and loop the idle animation when no other sequences are playing.

6. HEAD LOOK AND EYE CONTACT SCRIPT

A talking character is lifeless unless it looks at you once in a while. Bundled with EasyTalk is a script that allows your character to look at you, or another character, in a somewhat realistic way. It's random but seems to work well.

This is simply a head turn and eye-contact script. It does not actually turn your character around. That's up to your bone animation layer(s).

A small amount of intelligence is employed in the script. If you have a character typing on a keyboard, for example, he might look at you but his head will still be influenced by the original animation to a varying degree. And his eyes might turn back to his monitor while his head is turned to you.

If the character cannot turn his head far enough to look at you, he will do his best to look at you once in a while. If you are behind him, he will know it, and turn his head accordingly, but to a limited degree.

7. TERMINOLOGY

Here's a short list of some definitions that you should be familiar with:

Phoneme: Any of the perceptually distinct units of sound in a specified language that distinguish one word from another, for example p, b, d, and t in the words pad, pat, bad, and bat.

Viseme: A generic facial image, or mouth position that can be used to "describe" a particular sound. A popular misconception is that a viseme is a phoneme. That can be true if you map one mouth position for each phoneme in any given language but for the purposes of animation, this is seldomly used. For example, you can use a single viseme (mouth position) for a phoneme that represents a B, M, or P sound.

Blendshape: A blend shape, also known as a 'morph', is a shape interpolation. For our purposes, a blendshape is a "deformed" version of a select group vertices on a 3D mesh.

8. HELP AND RESOURCES

Visit Realtime VR on the web at http://www.realtimevr.com or email us at realtimevr@gmail.com.