

Synthesis in SynthWorks¹

Downloads at <http://www.linguistics.ucla.edu/people/grads/krisyu/teaching/lx104f10> under **Section materials** (Section 8) and Lab 4

1 Getting SynthWorks

- Available in CLICC lab (where we have section)
- Available in phonetics lab (Campbell 2101)
- SynthWorks Lite **for PC users**
 1. Go to <http://www.sciconrd.com/DownloadForm.aspx>
 2. Type in something in the form and click **Download**.
 3. You will then see a *Downloads and Upgrades* page.
 4. Click on the **Suite** under Windows to download demo versions of software including SynthWorks.
 5. **You will be able to play with SynthWorks, but not save your work!**

2 Review of Klatt formant synthesizer design

- Hybrid model: cascade plus parallel models for formant synthesis
 - Cascade: relative amplitudes of formant frequencies set based on frequencies of other formants
 - Parallel: amplitudes of formants independent
- Source-filter model
 - Sources:
 - * Voicing: on/off, loudness, f0, phonation
 - * Aspiration: on/off, loudness
 - * Frication: on/off, loudness, frequencies (for bursts, too)
 - Filters: cascade, parallel

3 The parameters

Source/voicing parameters

1. **f0** (x10), i.e. 1000 means 100 Hz, **not a source!**
2. **av** (voicing – cascade branch)
3. **ah** (aspiration – cascade branch)
4. **af** (frication – parallel branch)

At least one source must have a non-zero value for sound to be audible!

¹Based off of old notes from Roy Becker-Kristal. Thank you Roy!

Formant frequencies

- Formant frequencies **f5, f4, f3, f2, f1**
- In reverse order from highest to lowest frequency
- **Apply in both cascade and parallel branches**

Formant bandwidths (cascade)

- Formant bandwidths **b5, b4, b3, b2, b1**
- **Apply in only cascade branch**
- Defaults okay for vowels
- Bandwidths plus frequencies determine amplitudes (calculated in model, not set by you)

Parallel branch

- Formant amplitudes **a5f, a4f, a3f, a2f, a1f**, defaults 0/30, so you won't hear anything even if af on
- Formant bandwidths **b5f, b4f, b3f, b2f, b1f**, reasonable for consonants

4 Planning your synthesis

4.1 Exploring your model sound

1. Get a model sound file and open in Praat
2. Segment the sound file
 - (a) **Annotate > To TextGrid**
 - (b) All tier names: **seg**; leave point tiers blank
 - (c) Edit sound and textgrid file together
 - (d) Segment out acoustic events
3. Get file and segment durations
 - (a) File duration: in objects window, select Sound, **Query > Query time domain > Get selection length, _____ ms**
 - (b) Segment durations: in object window, select TextGrid, click **Down to Table**
 - (c) Set *Time decimal* to 3 (gives you 3 decimal places)
 - (d) You can save the table for Praat using **Write > Write to text file** and as a text file with **Write to table**.
4. Get formant track of the model sound
 - (a) In object window, select Sound, **Formants & LPC > To Formant (Burg)**
 - (b) Set *Time step* to 0.005 (5ms)
 - (c) Set *maximum formant* to 5000
 - (d) Click OK
 - (e) Click **Down to Table**
 - (f) Set *Time decimals* to 3 and click OK
 - (g) You can save the table for Praat using **Write > Write to text file** and as a text file with **Write to table**.

5. Get pitch track of the model sound

- (a) In object window, select Sound, **Periodicity** > **To Pitch**
- (b) Set *Time step* to 0.005 (5ms)
- (c) Click OK
- (d) Click **Down to PitchTier**
- (e) Click **Covert** > **Down to TableOfReal**, Click OK
- (f) Click **Convert** > **To Table**
- (g) You can leave *rowlabel* blank
- (h) You can save the table for Praat using **Write** > **Write to text file** and as a text file with **Write to table**.

4.2 Finding landmarks in your model sound

Segment	Start (ms)	End (ms)	Branch?	Formants?	Events?

Table 1: Model sound landmarks

Other notes?

5 Getting started with Synthworks

5.1 Opening Synthworks

- Open Synthworks, which is in the same folder as the other Linguistics programs
- You will see a giant spreadsheet, like an Excel spreadsheet
- The columns are timeframes (default: 200ms, 10-ms frames)
- The rows are acoustic parameters
- Don't worry! You won't actually need to set all of those numbers, just **anchors**. Then we will **interpolate** between the anchors.

5.2 Adjusting the length of the file

1. Before you start doing anything, you need to adjust the file length and parameters

2. Go to **Options > File**
3. Set *File Length* to the duration of the model sound file in ms
4. Set *Interval* to 5 ms (that's the time frame)
5. For the two other radio button options, you might want to leave them as is for now, but change them after you get comfortable using the program. Ideally, you want all your manual data to serve as anchors, and then all other values to be interpolated. But you also don't want to have a bunch of incorrect anchors to start with!

5.3 Selecting/organizing parameters

1. Go to **Options > Parameters**
2. Uncheck parameters we will not use. These include the **nasals** (no nasals in model sound) and **apv**, voicing amplitude in parallel branch (no voiced obstruents).
3. For each set of parameters, click on the **Color set** button and color code your parameters. **Don't use red! That color is reserved for anchor cells.**

5.4 Entering values, creating anchor cells and interpolating

1. To enter a value, just click on a cell and type in a number.
2. To manually change the cell into an anchor cell, double click on it. A box will ask you: "Change the cell into an anchor?" Click yes. Now you have a red anchor cell!
3. Enter a value in a cell about 10 frames to the right and change into an anchor cell
4. Now interpolate! Go to **Interpolation > Interpolate all**. This interpolates **linearly** between every pair of anchor cells in the spreadsheet!
5. To remove an anchor: double click on anchor cell and answer "yes".
6. To remove a set of anchors, select the range of cells containing those anchors and click on **Anchors > Remove from selection**
7. If you really want to go back to the drawing board, you can also get rid of every anchor you have with **Anchors > Remove All**.
8. What happens if you interpolate between two anchor cells of equal value? _____
9. What happens if you interpolate between two anchor cells of very different values? _____
10. How can you approximate curves with linear interpolation? _____

5.5 Parameter value ranges

Here are parameter ranges for parameters you will be using for your lab.

Parameter type	Parameter	Value range
F0	F0	500-5000 (50-500 Hz)
Sources	av, ah, af	20-87
Formant frequencies	f5, f4, f3	1200-4999
Formant frequencies	f2	550-3000
Formant frequencies	f1	200-1300
Formant bandwidths (cascade)	b5-b1	40-1000
Formant amplitudes (parallel)	a5f-a1f	20-87
Formant bandwidths (parallel)	b5f-b1f	40-1000

Table 2: Parameter value ranges for parameters used in lab

5.6 Loading the model file

- You can load the model file with **File > Open an audio file**.
- This will allow you to compare spectrograms/pitch/sound of the model to your synthesis.

5.7 Synthesizing/listening to your synthesis

- Click the play buttons (red triangles) to play your synthesis. You can play your current synthesis and the previous one. You can even play your model file.

5.8 Looking at your synthesis

- To see a spectrogram/pitch track of your synthesized utterance, click on the spectrogram icon (next to the FFT icon)
- Compare the spectrogram of your synthesis and that of the model often. You want them to look the same!

5.9 Saving your synthesis

- To save, go to **File > Save**. Save often!
- To save different versions of your synthesis, you can give them different names and select **Save As**
- To save an audio file of your synthesis, go to **File > Save as audio**.

5.10 Tips

- **Fricatives:** You need to type values into the **af** cells to activate the fricative noise source, and you also need to have at least one formant amplitude higher than 0. The **af** source turns on the parallel branch, while the **av** and **ah** sources turn on the cascade branch.
- **F0:** f_0 should never be 'zero'. Remember that f_0 is not a source. So if you set **av** to 0 (the voicing amplitude), the f_0 will be turned off. Don't forget that f_0 is in units of 0.1 Hz (so multiply by 10 when you enter values in Hz!)
- **Vowels:** leave bandwidths as they are to start with. You may be able to get quite a natural sounding synthesis with the default values.
- **Loudness:** To make something louder, increase the amplitude of the source, or increase the **gain** (one of the non-default parameters you can activate).
- **Smooth transitions:** if you hear clicks or pops, look for discontinuities along a parameter and smooth it out. Slow transitions from one segment to a next!
- **Start early!:** it only takes a couple hours to get the synthesis sounding intelligible. But many more to get it sounding natural!
- **Get help:** ask friends to listen for intelligibility, naturalness and closeness to model. At intermediate stages, you can talk to Professor Jun and me.