LPC Analysis and Resynthesis using Praat

A useful interactive tutorial is at:
http://www.kt.tu-cottbus.de/speech-analysis/tech.html

LPC analysis decomposes speech sounds into two parts:
1. a filter function consisting of LPC coefficients: based on the assumption that the source has been filtered through a single variable-cross-section tube (like the vocal tract);
2. a source function: which can either be:
   a. the original signal inverse filtered through the filter function
   b. a stylised version of (a) consisting of white noise (for unvoiced speech) or a pulse train (buzz, at the appropriate pitch for voiced speech).

The original signal can be recovered exactly by filtering 2a through 1.

An intelligible version of the original signal can be recovered by filtering 2b through 1. Transmitting 1 and 2b gives a substantial data reduction over the original signal, since the filter function only needs to specified every 10ms or so and consists of 10 or 12 real numbers, and the stylised source function needs specify (again every 10ms or so) the amplitude of the noise and the frequency and amplitude of the pulse train.

LPC coding was used extensively for low-bandwidth secure voice communication, but the quality can be unsatisfactory especially when the assumptions of the filter model are not met (as for nasal or fricative sounds) or if the voiced/voiceless decision is error prone.

**Praat steps**

Read in a waveform file such as wewere.aiff using Praat menu/ *Read from file*...

Use Info to check that the sampling rate is around 11kHz. If it is not, then use Convert/ *Resample*...
(The reason is so that the LPC analysis fits the spectrum with an appropriate number of poles which are spaced at about 1kHz interval in adult male speech).

Now do an LPC analysis on the wave using **Formants & LPC -/ To LPC (burg)**...

Select the LPC object and the original waveform **Filter (inverse)**

This operation creates a new Sound object which you can rename **LPCResidual**. The LPCResidual is the waveform which, when filtered through the 5-pole (11 coefficients) LPC filter, gives the original waveform. So it is an approximation to the speech source, and the LPC filter is an approximation to the vocal tract’s filtering.

Use Edit to look at the spectrograms of both the original sound and its LPCResidual. Notice that the formant structure has disappeared from the LPCResidual, but is still contains the harmonic structure (or noise for whispered speech) and the main amplitude variations. It also has some consonantal bursts (like /g/) which are not well-modelled by LPC.

You can reconstruct the original sound by selecting the LPC object and the LPCResidual sound and then clicking on Filter.

But to be more interesting, you can use any sound (at the same sampling frequency as the original speech, and at least as long) as the source function. You could use Shost3b_11025.aiff. To make the resulting sound louder use **Modify – Scale peak**...Have a look at the spectrogram of the resulting sound. When you listen to it you can hear the orchestra but a ghostly voice whispers the speech from within it.

You could experiment with the sounds of different instruments – swapping their source and formant structures. The process will work provided that the original sound has got the appropriate number of formants (poles). You can control the number of formants (**n**) by setting the Prediction order in the **To LPC (burg)** command to 2**n**. Single notes from musical instruments are available at: [http://theremin.music.uiowa.edu/MIS.html](http://theremin.music.uiowa.edu/MIS.html)
**Enharmonic sounds.** You can synthesise enharmonic speech, where the harmonics are sifted up by a constant number of Hz from harmonic values. This method does it in a fancy way by shifting the spectrum of the LPC residual rather than the raw waveform in order to ensure that the formant values are not shifted as well (remember the bassoon demo?).

1. flatten the pitch of the original using Change gender…
2. Do an LPC analysis on the flattened speech as above.
3. Select the LPCResidual from this analysis and run the following as a script (paste it into a new script created with Praat menu/ New script) after substituting the Hz shift you want in the first line:
   ```plaintext
   hzShift = 40
   Rename... newresidual
   To Spectrum... yes
   Copy... shiftedwav
   binwidth = Get bin width
   nbins = Get number of bins
   colshift = hzShift/binwidth
   Formula... if col=nbins then 0 else Spectrum_newresidual[row,col - colshift] fi
   To Sound
   ```
4. Treat this shiftedwave sound as the new LPCResidual, and filter it through the original LPC coefficients to produce the new speech.
   You will find that the new shifted speech has multiple pitches. Check that the script has worked properly by measuring some "harmonic" frequencies on a narrow-band spectrogram.

You can probably do something similar with instrument sounds.