We’ll be listening to

* Early synthesizer
* Synthesis by rule — researcher enters phoneme
* Complete text-to-speech system

Evolution Desche??has early synthesis was done (page4?)
- Different configuration  → Connected with Vcoders
<table>
<thead>
<tr>
<th>Tape Number</th>
<th>Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Voder</td>
<td>29.7</td>
</tr>
<tr>
<td>2. Pattern Playback</td>
<td>29.1</td>
</tr>
<tr>
<td>3. PAT</td>
<td>No Fig.</td>
</tr>
<tr>
<td>4. OVE</td>
<td>No Fig.</td>
</tr>
<tr>
<td>5. PAT2</td>
<td>29.8</td>
</tr>
<tr>
<td>6. OVE II</td>
<td>29.2</td>
</tr>
<tr>
<td>7. OVE II (holmes)</td>
<td></td>
</tr>
<tr>
<td>8. Holmes II Synthesis</td>
<td>29.3</td>
</tr>
<tr>
<td>9. Klatt (Male Fem.)</td>
<td>29.4</td>
</tr>
<tr>
<td>10. Dectalk</td>
<td></td>
</tr>
<tr>
<td>11. Davo</td>
<td>29.9</td>
</tr>
<tr>
<td>12. Flanagan</td>
<td>29.10 &amp; 29.11</td>
</tr>
<tr>
<td>13. Speach &amp; Spell</td>
<td>29.5</td>
</tr>
<tr>
<td>14. Multi pulse LPC</td>
<td></td>
</tr>
<tr>
<td>15. Pattern Playback</td>
<td>29.1</td>
</tr>
<tr>
<td>16. Kelly Gerstman</td>
<td>get Fig from previous 11.3</td>
</tr>
</tbody>
</table>
Synthesizers can be

- Channel vocoder, LPC or homomorphic
- Serial formants [each formant is a two-pole network]
- Parallel formants —
- Articulatory models
- Oddball arrangement pattern playback
Evolution

* Researcher pitches an utterance, creates a spectrogram.
* Researcher has a synthesizer model at his/her disposal.
* Researcher enter sequence of parameter values into model.
* Synthesizer “Speaks” and researcher adjusts sounds so utterance searches better, before this. We had the Voder where the instrument was “played” in real time by a skilled performer.
Speech Synthesis

Text to Phonemes

Graphemes to Phonemes

Phonemes to Parameters

Excitation

Speech Synthesizer

Speech

Synthesis by Rule

Prosodics?

Early Synthesizers

Synthesis by Rule

Complete text-to-speech.
Figure 29.7: Channel vocoder synthesizer.
Figure 29.8: Light Collector, mirror, Tone wheel, Spectrogram etc.
Figure 2.12: Spectrogram of “Greetings everybody” by announcer
Figure 29.8: Parallel formant synthesizer.
Figure 29.2: OVE II Speech Synthesizer of Gunnar Fant. Form [20]
Figure 29.4: The Klatt Synthesizer. From [35]. (cont.)
Figure 29.4: The Klatt Synthesizer. From [35].
Figure 29.9: DAVO (Dynamic analog of the vocal tract.) From []
Figure 29.10: Schematic of the vocal cord-vocal tract system.
Figure 29.11 : Circuit of an individual T-Section.
(a) Direct-Form Digital Filter with Variable “a” Coefficients

(b) Acoustic Tube with Variable Area Functions

Figure 29.5: Two configurations for all pole synthesizers based on LPC analysis. (cont.)
(c) All-Pole lattice Network with Variable “k” Parameters

Figure 29.5: Two configurations for all pole synthesizers based on LPC analysis.

a) shows a direct form implementation of the difference equation giving a synthesizer output as a weighted sum of its past values plus the excitation input. b) shows a model of the acoustic tube with variable cross-sectional area that could give rise to such a characteristic. c) shows an interpretation of this model that suggests a lattice form for the filter.
Figure 11.2: Two section digital wave guide.
Figure 29.6: All-Zero synthesizer based on depstral analysis.
Figure 29.12: Structure of Klatt Synthesizer.