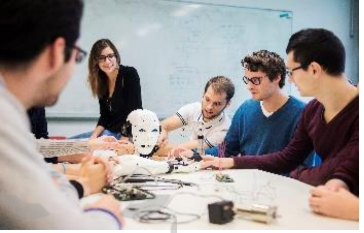
OBL-4101 Digital Signal Processing

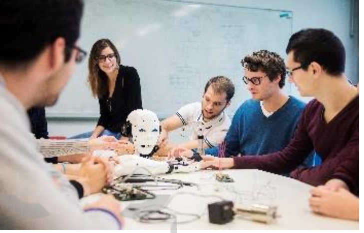
Project on voice coder

Responsible: N. Madaoui
English version : G. Lissorgues
emails : nadia.madaoui@esiee.fr and gaelle.lissorgues@esiee.fr



Presentation of the project

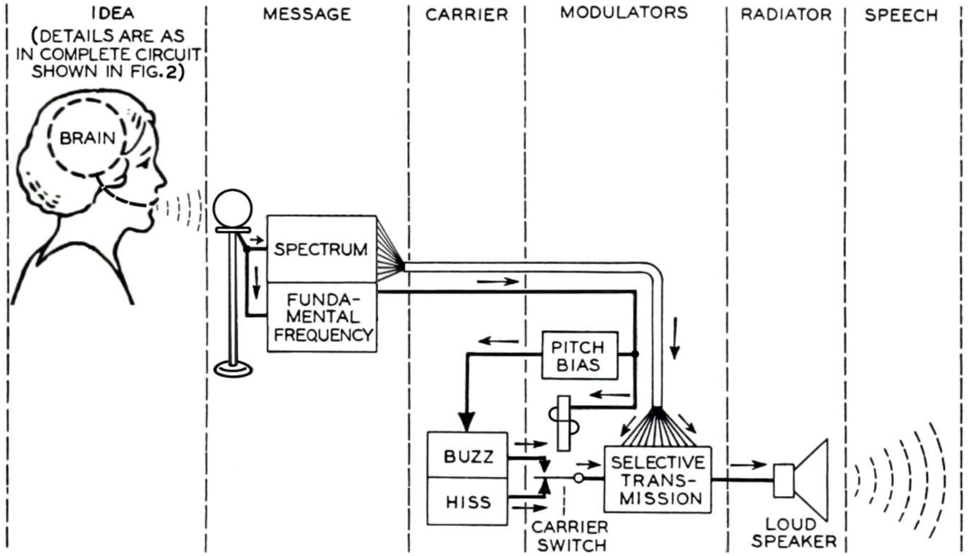
- **Context**
  - ✓ A voice coder also called vocoder is a system used to “code” voice artificially. The phase vocoder will mainly modify the phase of the signals, and “coding” here means “modification” and not data ou computer coding.
  - ✓ Remember that voice has frequencies between 60Hz to 1500Hz
  - ✓ Vocoder appeared during the second world war, then later in radio, movies, electronic music…



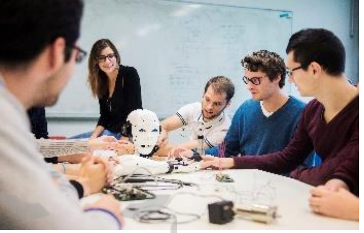
Presentation of the project

- **Context**
  - *Definition extracted from wikipedia*
  - A vocoder is a contraction of *voice* and *encoder* : it represents a category of voice codec that analyses and synthetises human voice signal
  - The vocoder was invented in 1938 by [Homer Dudley]() at [Bell Labs]()
  - By encrypting the control signals, voice transmission could be secured against interception. Its primary use was for secure radio communication.
  - Later, the vocoder was used in electronic musical instrument. In the 80ies, he first digital synthesizers were academic experiments in sound synthesis using digital computers.

Synclavier PSMT (1984)



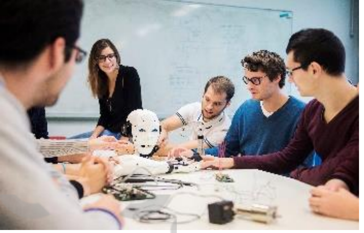
Homer Dudley (October 1940). "The Carrier Nature of Speech". Bell System Technical Journal, XIX(4);495-515. -- Fig.7 Schematic circuit of the vocoder (derived from Fig.8).jpg



Presentation of the project

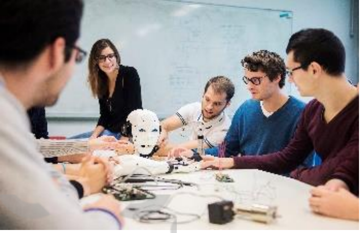
- **Phase vocoder**
  - ✓ 3 main effects will be studied:
    - modify the speed of the voice without changing its pitch. *(the sound of the voice is the same but the words will be pronounced more slowly or faster)*
    - modify the pitch of the voice without changing the speed. *(the pitch is related to the fundamental frequency of the voice and corresponds to the level i.e. high or low voice).*
    - apply an effect to transform the voice as if it comes from a robot ("robotisation")
  - ✓ Use given audio files or **create your own files to personalise** your work

And remember that: $y(t) = x(\alpha t) \overset{\text{FT}}{\Leftrightarrow} Y(f) = \frac{1}{|\alpha|} X(\frac{f}{\alpha})$ which explains compression/dilatation in both time and frequency domains



Presentation of the project

- **Work to be done**
  - ✓ The Matlab programs are given. As usual, the main parts are detailed. You should understand them and sometimes complete the code or modify it.
  - ✓ You will have to create 2 main functions:
    - The phase voice coder -> frequency interpolation
    - The robotisation of the voice
  - ✓ Minimal structure of the main program Vocoder.m:
    - Speed modification -> needs function PVoc.m
    - Pitch modification -> needs function PVoc.m + ré-échantillonnage (re-sampling)
    - Robotisation of the voice -> needs function Rob.m

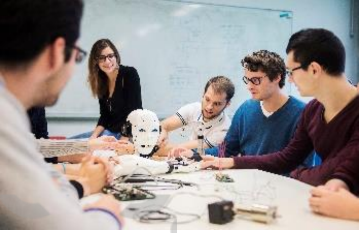


Presentation of the project

- **Work to be done**
  - ✓ Suggestions to go further: propose new ideas, new options, new signal processing treatments to improve the vocoder
  - ✓ Examples (not restricted): Audio effects can be classified by how works their signal processing
    - Basic Filtering — Lowpass, Highpass filters, Equalisers
    - Time Varying Filters — Wah-wah, Phaser
    - Delays — Vibrato, Flanger, Chorus, Echo
    - Modulators — Ring modulation, Tremolo, Vibrato
    - Non-linear Processing — Compression, Limiters, Distortion,
    - Exciters/Enhancers
    - Spacial Effects — Panning, Reverb, Surround Sound

CARDIFF UNIVERSITY PRIFYSGOL CAERDYDD

ESIEE PARIS



Presentation of the project

- **Work to be done**
  - ✓ Examples of basic Audio effects

Filters by definition **remove/attenuate** audio from the spectrum above or below some cut-off frequency.

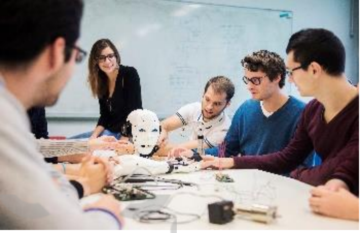
- For many audio applications this a little too restrictive

Equalisers, by contrast, **enhance/diminish** certain frequency bands whilst leaving others **unchanged**:

- Built using a series of *shelving* and *peak* filters
- First or second-order filters usually employed.

CARDIFF UNIVERSITY PRIFYSGOL CAERDYDD

ESIEE PARIS



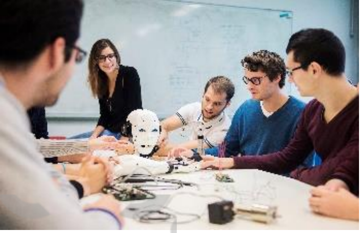
Presentation of the project

- **Work to be done**
  - ✓ Examples of other Audio effects : time varying effects

Some common effects are realised by simply time varying a filter in a couple of different ways:

Wah-wah — A bandpass filter with a time varying centre (resonant) frequency and a small bandwidth. Filtered signal mixed with direct signal.

Phasing — A notch filter, that can be realised as set of cascading IIR filters, again mixed with direct signal.



Presentation of the project

- **Work to be done**
  - ✓ Examples of other Audio effects : reverberation

Reverberation (**reverb** for short) is probably one of the most heavily used effects in music.

Reverberation is the result of the many reflections of a sound that occur in a room.

- From any sound source, say a speaker of your stereo, there is a direct path that the sounds covers to reach our ears.
- Sound waves can also take a slightly longer path by reflecting off a wall or the ceiling, before arriving at your ears.

Echo — implies a distinct, delayed version of a sound,

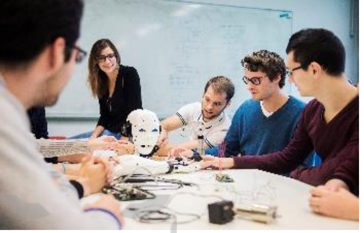
- *E.g.* as you would hear with a delay more than one or two-tenths of a second.

Reverb — each delayed sound wave arrives in such a short period of time that we do not perceive each reflection as a copy of the original sound.

- Even though we can't discern every reflection, we still hear the effect that the entire series of reflections has.

CARDIFF UNIVERSITY PRIFYSGOL CAERDYḎ

ESIEE PARIS



Presentation of the project

- **Work to be done**
  - ✓ Example of robotisation effect : the ring modulation

Ring modulation (RM) is where the audio *modulator* signal, $x(n)$ is multiplied by a sine wave, $m(n)$, with a *carrier* frequency, f_c.

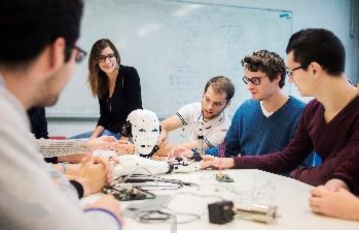
- This is very simple to implement digitally:

$$y(n) = x(n).m(n)$$

- Although audible result is easy to comprehend for simple signals things get more complicated for signals having numerous partials
- If the modulator is also a sine wave with frequency, f_x then one hears the sum and difference frequencies: $f_c + f_x$ and $f_c - f_x$, for example.
- When the input is *periodic* with at a fundamental frequency, f_0, then a spectrum with amplitude lines at frequencies $|kf_0 \pm f_c|$
- Used to create robotic speech effects on old sci-fi movies and can create some odd almost non-musical effects if not used with care.

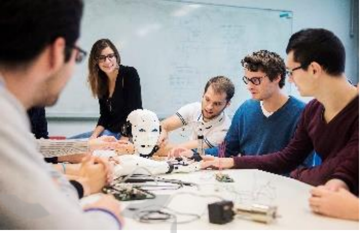
CARDIFF UNIVERSITY PRIFYSGOL CAERDYDD

ESIEE PARIS



Presentation of the project

- **Work to be done**
  - ✓ Search also on bibliography references on the topic
  - ✓ Look at Matlab community
  - ✓ And have fun !



Presentation of the project

- **Additionnal notion (necessary for the project)**
  - ✓ The Short Term Fourier Transform (STFT) – link with TP1 and spectrograms
  - ✓ You will need to split the voice signal into frames of 30ms (approximate usual estimation in speech processing) where you will consider the signal stationnary
  - ✓ The Short-Time Fourier Transform (STFT) (or short-term Fourier transform) is a powerful general-purpose tool for audio signal processing. It defines a particularly useful class of time-frequency distributions which specify complex amplitude versus time and frequency for any signal.
  - ✓ See theory details here: https://www.dsprelated.com/freebooks/sasp/Short_Time_Fourier_Transform.html

+ a nice video explaining the concept: https://www.youtube.com/watch?v=8nZrgJjl3wc

ESIEE PARIS