A prominent linguistic segment is a segment which is pronounced to stand out of its context.

Acoustic Prominence has strong visual correlates: facial gestures are strong indicators of prominence, e.g. they can indicate prominence when speech is noisy.

Questions:
- Can we detect prominence in real-time from the speech signal so to produce its facial correlates in visual speech synchrony applications?
- Does acoustic prominence aids speech perception? can only facial gestures aid speech perception instead of their acoustic correlates?
- When and how shall we animate the face to trigger the perception of prominence? and would this influence the human-likeness and the face reading behavior of subjects?

**SynFace**, a speech driven facial animation system, uses only the phonetic sequence from speech. For the use in SynFace: detect relative prominence in real-time over phonetic segments:

**Method:**
Estimate vowel prominence using prosodic parameters (vowel duration, pitch, loudness)
Over a window, place a gesture over the vowel which has the highest estimated prominence

If acoustic prominence enhances speech comprehension, can gestures synchronized with it have the same effect?

**Setup:**
Degrade the acoustic signal using a noise excited vocoder.
Support the acoustic signal by a lip-synchronized talking head
Measure speech intelligibility change when different gestures are applied to the face at different controlled timing.

12 Subjects are used in the experiment, the subject listens to a vocoded, semantically complete sentence and writes down what they perceive.
6 different visual stimulus were used.

The percentage of time spent on the eyes, mouth and the rest of the face in the two conditions.

**Gaze behavior**
Face without gestures
Face with Gestures

**The distribution (Boxplot) of the recognition rate per sentence over the different visual variants**

**Questionnaires Results**
The percentage of time spent on the eyes, mouth and the rest of the face in the two conditions.