Quiz #1

1. Consider the portuguese word “javali”. If we toggle in the phonetic sequence only the voiced/unvoiced characteristic of the 2 first consonants which would be the phonological sequence obtained:

a) xabali.
b) zavali.
c) fazali.
d) xafali.

2. The oral production of vowels is associated to vocal tract ressonances. The respective frequencies are named formants. Regarding these we can say the following:

a) They are harmonics.
b) Depend on the articulation.
c) Depend on the tone.
d) Don’t depend on the speed of sound in air .

3. In the source-filter model of speech production there is place for:

a) Excitation signal with pulses, noise or a mixture of both types.
b) The transfer function of the vocal tract.
c) The mouth radiation correction.
d) All the choices are correct.

4. In a speech signal zone with very low zero crossing rate values there can be:

a) Silence with offset.
b) Voiced signals.
c) Silence.
d) Unvoiced signals.

5. In the determination of f0 for a speech signal segment the auto-correlation and peak detection technique was used. For that the following should have been used:

a) A window with duration much larger than the average value of the period to be measured and equal step size.
b) A window with duration slightly larger than the average value of the period to be measured and step size maller than the smaller period to be measured.
c) A window with duration much larger than the average value of the period to be measured and step size equal to the period to be measured.
d) A window with duration slightly larger than the average value of the period to be measured and equal step size.

6. The spectrogram of a speech signal is an example of a short-term time-shifting analysis. A time window is used and the result of the transform of each segment shows the short-time spectrum. Among the free variables to adjust the visualization of the signal’s characteristics we can choose the window type, the duration and the superposition:

a) The time window duration strongly influences the resulting frequency resolution, the one obtained with a larger window being the better.
b) The step size influences the resulting frequency resolution.
c) The used time-window duration doesn’t influence significantly the resulting frequency resolution.
d) The visualization of formants is done preferably with longer duration windows.

7. In the speech signal prosody consists of:

a) Intonation.
b) choice a) plus segmental durations.
c) choice b) plus signal/noise ratio.
d) choice c) plus energy.

8. Knowing that formant frequencies of the speech signal indicate certain characteristics of the vocal tract that has produced it, these frequencies can be determined:

a) By determination of peaks of the spectrum of the speech signal determined with a window with adequate length.
b) Directly from the amplitude of the voice signal.
c) By means of the LPC coefficients of the glottal excitation.
d) All other choices are wrong.

9. The LPC analysis or coding consists in the determination of a set of N coefficients of a descriptive function for a speech production system:

a) The coefficients are the roots of the system function.
b) The normal equations system roots are the coefficients.
c) The N+1 normal equations are symmetric.
d) To build the normal equations system with N equations it is necessary to calculate (N+1)*N different values.

10. The fundamental frequency of a speech signal can be calculated through:

a) Autocorrelation, AMDF, peaks and valleys and LPC coefficients.
b) The items in a) plus covariance.
c) Detection of maxima or minima of a certain function calculated from the sinal of interest.
d) All other choices are wrong.

11. Text pre-processing contributes to reduce various text deficiencies and perform conversions. Give some significant and illustrative examples of these functions and discuss the objectives.

12. Explain the function of the glottis during the production of voiced speech sounds.