1. The biometric modality is not:

☐ fingerprint,

☐ voice,

☐ password or pin (personal identification number),

☐ iris.

---

2. Determine the proper sequence of stages/phases in a general recognition process. (1 – first, 2 – second, 3 – third, 4 – last).

a) ___ recognition,

b) ___ data acquisition,

c) ___ data registration,

d) ___ pre-processing.

---

3. Multi-level user identification uses:

☐ user's voice,

☐ user's face,

☐ credentials as a username, password or PIN,

☐ combination of user's voice, face and credentials.

Erasmus+
This project has been funded with support from the European Commission.
This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

4. Match right definition to each kind of speaker identification system below :

Text independent systems	use extracted acoustic information e.g. formant frequencies, spectra, etc.
Text dependent systems	use precise phrases or passwords
Systems utilizing acoustic information	use extracted prosodic information e.g. speech dynamic, stress, pauses, etc.
Systems utilizing prosodic information	do not use precise phrases or passwords.

---

5. Determine the right sequence of operations/ processes in the automatic speech recognition process.
(1 – first, 2 – second, 3 – last).

a) ___ dictionary search,

b) ___ digital speech signal,

c) ___ text sequence.

---

6. The most successful methods for extracting speech feature are:

☐ Hidden Markov models,

☐ Mel frequency cepstral coefficients,

☐ Perceptual linear prediction,

☐ Discrete cosine transform.

Erasmus+
This project has been funded with support from the European Commission.
This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

7. Match each automatic speech recognition (ASR) system with its proper input definition.

ASR recognizing isolated words require	the input to be natural speech without grammatical restrictions.
Dictation systems require	the input to be a single word from a dictionary.
ASR recognizing fluent speech require	the input to be fluent speech with some grammatical restrictions.
ASR recognizing natural speech require	the input to be a sequence of words with sufficient pauses separating adjacent words.

---

8. Classify given spectral events/modifications to perceivable and non-perceivable by humans.

Perceivable	Non-perceivable

A – Number of formant frequencies,
B – Frequencies laying under the first formant frequency,
C – Location of formant frequencies,
D – Width of formant frequencies,
E – Overall tilt of the spectra,
F – Narrow band stop filtering.

Erasmus+
This project has been funded with support from the European Commission.
This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

9. The most significant speech recognition methods are:

☐ Hidden Markov models,

☐ Dynamic time warping,

☐ Perceptual linear prediction,

☐ Discrete cosine transform.

---

10. The main purpose of dynamic time warping in speech recognition domain is:

☐ Feature extraction.

☐ Comparison of two sequences of speech features that differ in time duration and to calculate their similarity.

☐ Calculation of spectral coefficients.

Erasmus+
This project has been funded with support from the European Commission.
This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.