Speech prosody (rhythm, intonation, duration) constitutes the supra-segmental framework of both speech production and speech perception.

On the production side, it determines the phonetic realisations of phoneme durations, stress patterns, syntagmatic grouping of morphemes (dependency relations), discourse types (statements, requests) and emotional meanings.

On the perception side, prosody disambiguates the speech signal by facilitating the segmentation of an acoustic continuum into meaningful linguistic units. The structural features of music also include a supra-segmental level (melody, tempo, harmony) and a segmental level (individual sounds).

While rhythm – be it in speech or music- is an acoustic phenomenon, its perception involves the listeners’ motor reactions. The learning of rhythm and intonation is thus independent of the utterance’s lexical meaning. Cognitive and neurological mechanisms that subserve the acquisition of prosody and phonology will be presented within the framework of Piaget’s genetic epistemology and Liberman’s motor theory of speech perception.

In spite of substantial evidence that prosody organises both speech production and speech perception, foreign language teaching methods often neglect it. To demonstrate how the acquisition of prosody can and should precede the learning of lexical meanings, we elaborated a foreign language teaching methodology which includes

1. Conversion of spoken dialogues into vocal hums and musical instruments,
2. Visualisation melodic contours and their gestural imitation
3. Variations of emotional tones of dialogues (sound morphing).

Our method confirms that the acquisition of musical aspects of speech is independent of lexical meanings and that audio-motor training enhances speech perception and phonological processing of spoken language.