

SECTION G — PHYSICS

G10 MUSICAL INSTRUMENTS; ACOUSTICS

G10L SPEECH ANALYSIS OR SYNTHESIS; SPEECH RECOGNITION; AUDIO ANALYSIS OR PROCESSING [4]

Note(s) [2010.01]

This subclass does not cover :

- devices for the storage of speech or audio signals, which are covered by subclasses G11B and G11C;
- encoding of compressed speech signals for transmission or storage, which is covered by group H03M 7/30.

11/00	Determination or detection of speech or audio characteristics not restricted to a single one of groups G10L 15/00-G10L 21/00 [7]	15/24	• Speech recognition using non-acoustical features, e.g. position of the lips [7]
11/02	• Detection of presence or absence of speech signals [7]	15/26	• Speech to text systems (G10L 15/08 takes precedence) [7]
11/04	• Pitch determination of speech signals [7]	15/28	• Constructional details of speech recognition systems [7]
11/06	• Discriminating between voiced and unvoiced parts of speech signals (G10L 11/04 takes precedence) [7]	17/00	Speaker identification or verification [7]
13/00	Speech synthesis; Text to speech systems [7]	19/00	Speech or audio signal analysis-synthesis techniques for redundancy reduction, e.g. in vocoders; Coding or decoding of speech or audio signals, e.g. for compression or expansion, source-filter models or psychoacoustic analysis [7]
13/02	• Methods for producing synthetic speech; Speech synthesisers [7]	19/02	• using spectral analysis, e.g. transform vocoders or subband vocoders [7]
13/04	• Details of speech synthesis systems, e.g. synthesiser structure or memory management [7]	19/04	• using predictive techniques [7]
13/06	• Elementary speech units used in speech synthesisers; Concatenation rules [7]	19/06	• Determination or coding of the spectral characteristics, e.g. of the short term prediction coefficients [7]
13/08	• Text analysis or generation of parameters for speech synthesis out of text, e.g. grapheme to phoneme translation, prosody generation or stress or intonation determination [7]	19/08	• Determination or coding of the excitation function; Determination or coding of the long-term prediction parameters [7]
15/00	Speech recognition (G10L 17/00 takes precedence) [7]	19/10	• Determination or coding of a multipulse excitation [7]
15/02	• Feature extraction for speech recognition; Selection of recognition unit [7]	19/12	• Determination or coding of a code excitation, e.g. in code excited linear prediction [CELP] vocoders [7]
15/04	• Segmentation or word limit detection [7]	19/14	• Details not provided for in groups G10L 19/06-G10L 19/12, e.g. gain coding, post filtering design or vocoder structure [7]
15/06	• Creation of reference templates; Training of speech recognition systems, e.g. adaptation to the characteristics of the speaker's voice (G10L 15/14 takes precedence) [7]	21/00	Processing of the speech signal to produce another audible or non-audible signal, e.g. visual or tactile, in order to modify its quality or its intelligibility (G10L 19/00 takes precedence) [7]
15/08	• Speech classification or search [7]	21/02	• Speech enhancement, e.g. noise reduction or echo cancellation (reducing echo effects in line transmission systems H04B 3/20; echo suppression in hand-free telephones H04M 9/08) [7]
15/10	• using distance or distortion measures between unknown speech and reference templates [7]	21/04	• Time compression or expansion [7]
15/12	• using dynamic programming techniques, e.g. Dynamic Time Warping [DTW] [7]	21/06	• Transformation of speech into a non-audible representation, e.g. speech visualisation or speech processing for tactile aids (G10L 15/26 takes precedence) [7]
15/14	• using statistical models, e.g. Hidden Markov Models [HMM] (G10L 15/18 takes precedence) [7]	23/00	Speech analysis not provided for in other groups of this subclass [2009.01]
15/16	• using artificial neural networks [7]		
15/18	• using natural language modelling [7]		
15/20	• Speech recognition techniques specially adapted for robustness in adverse environments, e.g. in noise or of stress induced speech (G10L 21/02 takes precedence) [7]		
15/22	• Procedures used during a speech recognition process, e.g. man-machine dialog [7]		