CLASS 704, DATA PROCESSING: SPEECH SIGNAL PROCESSING, LINGUISTICS, LANGUAGE TRANSLATION, AND AUDIO COMPRESSION/DECOMPRESSION

SECTION I - CLASS DEFINITION

This is the generic class for apparatus and corresponding methods for constructing, analyzing, and modifying units of human language by data processing, in which there is a significant change in the data.

This class also provides for systems or methods that process speech signals for storage, transmission, recognition, or synthesis of speech.

This class also provides for systems or methods for bandwidth compression or expansion of an audio signal, or for time compression or expansion of an audio signal.

Class 704 is structured into three main divisions:

A. Linguistics.

B. Speech Signal Processing.

C. Audio Compression.

See Subclass References to the Current Class, below, for the subclasses located within each of these three main divisions.

SECTION II - LINES WITH OTHER CLASSES AND WITHIN THIS CLASS

A. LINGUISTICS

1. This class does not include subject matter wherein significant details of the modification or construction of documents are claimed. (See Class ?0? in the Search Class notes below in References to Other Classes, regarding Document Processing).

2. This class does not include subject matter directed to significant details of teaching languages. (See Class 434 in the Search Class notes in References to Other Classes, below).

3. This class does not include subject matter directed to significant details of the construction, analysis or modification of computer languages. (See Class 717 in the Search Class notes in References to Other Classes, below).

B. IMAGE ANALYSIS

1. This class does not include subject matter wherein significant image analysis is performed and speech signal processing is nominally claimed (see Class 382 in the Search Class notes in References to Other Classes, below).

2. This class includes subject matter directed to speech signal processing disclosed or claimed in plural diverse arts such as image analysis (classified, per se, in Class 382).

C. AUDIO SIGNAL PROCESSING

1. This class does not include subject matter wherein nominal bandwidth or time modifications are performed for other audio processing defined in Classes 381 or 84 (see Search Class notes below in References to Other Classes). Examples of subject matter not included are: Stereo, sound effects, hearing aids, input and output transducers, and musical instruments.

2. This class includes audio signal processing wherein significant processing is performed to modify the signal"s bandwidth or time characteristics for compression or expansion of the signal.

D. COMMUNICATIONS

1. This class does not include subject matter wherein significant details of a distinct communications system or telephone link is performed and speech signal processing is nominally claimed (see Classes 340, 370, 375, 379, 455 in the Search Class notes below in References to Other Classes.).

2. This class includes subject matter directed to speech signal processing disclosed or claimed in plural diverse arts such as various types of communication systems.

E. APPLICATIONS

1. This class does not include subject matter wherein significant details of application systems are performed and speech signal processing is nominally claimed.

2. This class includes subject matter directed to speech signal processing disclosed or claimed in plural diverse arts to include electrical and mechanical systems. Examples would include systems controlled by speech recognition, systems which create specific displays of speech data, systems for editing speech data and other-

wise unrelated systems which incorporate <u>speech</u> signal processing details such as placing a <u>speech</u> synthesizer into novelty items.

SECTION III - SUBCLASS REFERENCES TO THE CURRENT CLASS

SEE OR SEARCH THIS CLASS, SUBCLASS:
1+,   for linguistics.
100+,   for speech signal processing.
500+,   for audio compression.

SECTION IV - REFERENCES TO OTHER CLASSES

SEE OR SEARCH CLASS:
84,   Music, subclasses 1+ for instruments used in producing music to include (a) electrical music instruments, (b) automatic instruments, and (c) hand-played instruments. Automatic and hand-played instruments are divided into four groups: stringed, wind, rigid vibrators, and membranes. This class also includes some accessory devices generally recognized as belonging to the art or industry.

181,   Acoustics, various subclasses, for mechanically transmitting, amplifying and ascertaining the direction of sound and for mechanically muffling or filtering sound.

340,   Communications: Electrical, subclasses 1.1 through 16.1 for controlling one or more devices to obtain a plurality of results by transmission of a designated one of plural distinctive control signals over a smaller number of communication lines or channels.

341,   Coded Data Generation or Conversion, various subclasses for electrical pulse and digit code converters (e.g., systems for originating or emitting a coded set of discrete signals or translating one code into another code wherein the meaning of the data remains the same but the formats may differ).

345,   Computer Graphics Processing and Selective Visual Display Systems, various subclasses for the selective control of two or more light generating or light controlling display elements in accordance with a received image signal, and subclasses 1.1 through 3.4 for visual display systems with selective electrical control including display memory organization and structure for storing image data and manipulating image data between a display memory and display device.

360,   Dynamic Magnetic Information Storage or Retrieval, which is an integral part of Class 369 following subclass 18, for record carriers and systems wherein information is stored and retrieved by interaction with a medium and there is relative motion between a medium and a transducer, for example, magnetic disk drive devices, and control thereof, per se.

365,   Static Information Storage and Retrieval, various subclasses for addressable static singular storage elements or plural singular storage elements of the same type (i.e., the internal elements of memory, per se).

369,   Dynamic Information Storage or Retrieval, various subclasses for record carriers and systems wherein information is stored and retrieved by interaction with a medium and there is relative motion between a medium and a transducer.

370,   Multiplex Communications, for the simultaneous transmission of two or more signals over a common medium, particularly subclasses 58.1+ for time division multiplex (TDM) switching, subclasses 85.1+ for time division bus transmission, and subclasses 91+ for asynchronous TDM communications including addressing.

375,   Pulse or Digital Communications, various subclasses for generic pulse or digital communication systems and synchronization of clocking signals from input data.

377,   Electrical Pulse Counters, Pulse Dividers, and Shift Registers: Circuits and Systems, various subclasses for generic circuits for pulse counting.

379,   Telephonic Communications, various subclasses for two-way electrical communication of intelligible audio information of arbitrary content over a link including an electrical conductor.

380,   Cryptography, appropriate subclasses for cryptographic electric signal modification.

381,   Electrical Audio Signal Processing Systems and Devices, various subclasses for wired one-way audio systems, per se.

382,   Image Analysis, various subclasses for operations performed on image data with the aim of measuring a characteristic of an image, detecting variations, detecting structures, or transforming the image data, and for procedures for analyzing and categorizing patterns present in image data.

434,   Education and Demonstration, subclasses 112+ for communication aids for the handicapped, subclasses 156+ for education and demonstration of <u>language</u>, subclasses 322+ for question or problem eliciting response.

455,   Telecommunications, appropriate subclasses for modulated carrier wave communication, per se, and subclass 26.1 for subject matter which blocks access to a signal source or otherwise limits usage of modulated carrier equipment.

700,   Data Processing: Generic Control Systems or Specific Applications, subclasses 1 through 89 for data processing generic control systems, subclasses 90-306 for applications of computers in various environments.

702,   Data Processing: Measuring, Calibrating, or Testing, appropriate subclasses for the application of computer data processing in measuring, calibrating, or testing.

708,   Electrical Computers: Arithmetic Processing and Calculating, subclasses 1+ for hybrid computers, subclasses 100+ for calculators, digital signal processing and arithmetical processing, per se, subclasses 300+ for digital filters, and subclasses 800+ for electric analog computers.

713,   Electrical Computers and Digital Processing Systems: Support, subclass 187 and 188 for software program protection or computer virus detection in combination with data encryption.

714,   Error Detection/Correction and Fault Detection/Recovery, various subclasses for generic electrical pulse or pulse coded data error detection and correction.

715,   Data Processing: Presentation Processing of Document, Operator Interface Processing, and Screen Saver Display Processing, subclasses 243 through 272 for document processing including layout, editing, and spell-checking.

717,   Data Processing: Software Development, Installation, and Management, appropriate subclasses for significant details of the construction, analysis, or modification of computer languages.

SECTION V - GLOSSARY

The terms below have been defined for purposes of classification in this class and are shown in <u>underlined type</u> when used in the class and subclass definitions. When these terms are not underlined in the definitions, the meaning is not restricted to the glossary definitions below.

CORRELATION

A statistical measurement of the interdependence or association between two variables that are quantitative or qualitative in nature. A typical calculation would be performed by multiplying a signal by either another signal (cross-correlation) or by a delayed version of itself (autocorrelation).

DISTANCE

A statistical measurement for comparing elements defined by variables or vectors using scalar or vector subtraction of those elements. Examples: <u>distance</u>=a-b, |a-b|, (a-b).5 or two vectors may be treated as objects such that the straight line distance is measured between them.

EXCITATION

Stimulation of the vocal tract by vibratory action of the vocal cords or by a turbulent air flow. In a digital system, the vocal tract is typically modelled with a filter and <u>excitation</u> of the filter is performed using time representations of <u>pitch</u> (<u>voiced</u> <u>excitation</u>) and <u>noise</u> (<u>unvoiced</u> <u>excitation</u>).

LANGUAGE

A systematic means of communicating ideas or feelings by the use of conventionalized sounds, gestures, or marks having understood meanings.

LINGUISTICS

The study of human <u>speech</u> including the units, nature, structure, and modification of <u>language</u>.

MASKING

1. The interference with the perception of one sound (the signal) with another sound (the masker). 2. The number of decibels by which a masking sound will raise (or change) a listener"s threshold of audibility of other sounds.

CRITICAL BANDWIDTHS

Bandwidths of the hearing process, as measured by the masking effect of a white, random noise in which a person detects a pure tone.

BARK SPECTRUM

The width of one critical band.

MEL

A subjective measure of pitch based upon a signal of 1000 Hz. being defined as "1000 mels" where a perceived frequency twice as high is defined as 2000 mels and half as high as 500 mels.

NOISE

Any sound which is undesirable and interferes with one"s hearing or with a system"s analysis of desired sound.

PHON

The loudness level of any other sound based upon the SPL (sound pressure level measured in decibels) of a 1 kHz tone. For example, if we judge a certain waveform to sound as loud as a 1 kHz tone at 70 dB, then this waveform has a loudness level of 70 phons.

PITCH

The measurable frequency or period at which the glottis vibrates.

SIMILARITY

A statistical measurement which is inversely proportional to distance. For example, if two patterns are compared yielding a small distance, then the patterns would exhibit a large (or high degree of) similarity.

SONE

A measure of loudness as a function of frequency and sound pressure. A pure tone of 1 kHz. at 40 db above a normal listener"s threshold produces a loudness of 1 sone.

SPEECH

The communication or expression of thoughts in spoken words.

UNVOICED

Speech sounds produced by a turbulent flow of air created at some point of stricture in the vocal tract and usually lacking pitch.

VOICED

Speech sounds produced by vibratory action of the vocal cords and usually having pitch.

SUBCLASSES

1 LINGUISTICS:
This subclass is indented under the class definition. Subject matter including means or steps for constructing a word, a phrase, or a sentence in a language.

SEE OR SEARCH CLASS:
434,   Education and Demonstration, subclasses 156+ for demonstration and education in linguistics.

2 Translation machine:
This subclass is indented under subclass 1. Subject matter wherein a language (i.e., source language) stored in a memory means is translated into another language (i.e., target language).

SEE OR SEARCH THIS CLASS, SUBCLASS:
9,   for translation machines with significant natural language processing.

SEE OR SEARCH CLASS:
358,   Facsimile and Static Presentation Processing, subclass 403 for document filing and retrieval system.
716,   Computer-Aided Design and Analysis of Circuits and Semiconductor Masks, subclasses 103 through 105 for translation of computer program in designing and analyzing circuits and semiconductor mask.
717,   Data Processing: Software Development, Installation, and Management, subclasses 136 through 161 for software program code translator or compiler in software development.

3 Having particular Input/Output device:
This subclass is indented under subclass 2. Subject matter wherein the translation machine includes a means for reading into the memory means a language, for pronouncing the translated language or a particular user interface.

(1)   Note. Examples of such devices include an optical scanner or voice synthesizer.

4   Based on phrase, clause, or idiom:
This subclass is indented under subclass 2. Subject matter wherein the translation machine translates a series of words that form a syntactical unit.

5   For partial translation:
This subclass is indented under subclass 2. Subject matter wherein the translation machine includes a means for providing translation for a specified portion of a sentence or a clause.

6   Punctuation:
This subclass is indented under subclass 2. Subject matter wherein the translation machine translates a compound word formed by hyphenation or sentences with quotation marks, colons, semicolons, or parentheses.

7   Storage or retrieval of data:
This subclass is indented under subclass 2. Subject matter including a means for assigning storage locations or accessing addresses to the memory means.

SEE OR SEARCH CLASS:
707,   Data Processing: Database, Data Mining, and File Management or Data Structures subclasses 736   through 757 for preparing data for information retrieval including clustering, generating an index, ranking, scoring and weighting records, latent semantic indexing, subclass 760 for translating queries between languages and 794 for semantic network data structures.

8   Multilingual or national language support:
This subclass is indented under subclass 1. Subject matter including means or steps to adapt to, process, or support plural languages in systems or in software (i.e., providing language identifiers on files or providing screen prompts in a selected language), or to support the conventions or peculiarities of various national languages (i.e., alphabetical ordering, date or currency indications).

SEE OR SEARCH THIS CLASS, SUBCLASS:
200+,   for details of translation between multiple languages.

SEE OR SEARCH CLASS:
715,   Data Processing: Presentation Processing of Document, Operator Interface Processing, and Screen Saver Display Processing, subclasses 264 through 265 for composing or editing multiple languages in a document and subclass 866 for customization or edition of operator interfaces.

9   Natural language:
This subclass is indented under subclass 1. Subject matter includes a means for applying grammatical rules or other analyses (e.g., morphemic, syntax, semantic, etc.) to define the true meaning of a sentence or phrase.

(1)   Note. When words are undefined in the dictionary of a natural language, the grammatical rules or other analyses are applied in order to determine the true meaning of a sentence or a phrase.

SEE OR SEARCH CLASS:
707,   Data Processing: Database, Data Mining, and File Management or Data Structures, subclasses 736   through 757 for preparing data for information retrieval including clustering, generating an index, ranking, scoring and weighting records, latent semantic indexing, subclass 760 for translating queries between languages and 794 for semantic network data structures.

10   Dictionary building, modification, or prioritization:
This subclass is indented under subclass 1. Subject matter including a construction, a change, or an orderly arrangement of dictionary, thesauri, or the like.

SEE OR SEARCH THIS CLASS, SUBCLASS:
9,   for mere use in natural language processing.
200+,   for mere use in translation.

SEE OR SEARCH CLASS:
707,   Data Processing: Database, Data Mining, and File Management or Data Structures, subclasses 736 through 757 for preparing data for information retrieval including clustering, generating an index, ranking, scoring, weighting records and database details of dictionaries.

715,   Data Processing: Presentation Processing of Document, Operator Interface Processing, and Screen Saver Display Processing, subclasses 259 through 260 for mere use of a dictionary in editing or composition of a document.

200   SPEECH SIGNAL PROCESSING:
This subclass is indented under the class definition. Subject matter wherein the system performs operations or functions on signals which represent <u>speech</u>.

SEE OR SEARCH THIS CLASS, SUBCLASS:
500+,   for audio (other than <u>speech</u>) signal bandwidth compression or expansion.

SEE OR SEARCH CLASS:
379,   Telephonic Communications, appropriate subclasses for <u>speech</u> signal processing in a telephone system or device.

200.1   Psychoacoustic
This subclass is indented under subclass 200. Subject matter wherein an operation on the signal is based upon the masking behavior of the human auditory system.

(1)   Note.   The calculation of masking thresholds based upon incoming analysis of audio is the basis of psychoacoustic compression because the frequency with the highest local amplitude will tend to mask (make inaudible) nearby frequencies below the threshold.

(2)   Note. MPEG (Motion Picture Experts Group) sets international standards such as MPEG 1, level 3 (commonly called MP3) for psychoacoustic coding to achieve audio compression of up to 10:1.

Typical coders work on a 16-bit PCM audio signal, which is the typical CD quality standard.

(3)   Note. Only white noise in a bandwidth centered about a tone and less than or equal to the critical bandwidth contributes to the masking effect.   Critical bands are generally considered a set of filters or channels tuned to different center frequencies having a bandwidth of less than a third of an octave.

(4)   Note. A plot of frequency versus pitch in mels is similar in shape to the plot of frequency versus the position of auditory-nerve patches on the basilar membrane.   This is evidence that human judgment of pitch is based upon the point of excitation along the basilar membrane in the ear.

SEE OR SEARCH CLASS:
382,   Image Analysis, subclass 239 for adaptive coding used in MPEG, JPEG & motion JPEG images.

201   For storage or transmission:
This subclass is indented under subclass 200. Subject matter wherein the <u>speech</u>, which may be in coded or reduced formats, is stored or transmitted.

202   Neural networks:
This subclass is indented under subclass 201. Subject matter wherein coding is performed using parallel distributed processing elements constructed in hardware or simulated in software.

SEE OR SEARCH THIS CLASS, SUBCLASS:
259,   for neural networks which decode a coded speech signal.

203   Transformations:
This subclass is indented under subclass 201. Subject matter wherein the <u>speech</u> is encoded using a specific mathematical function (e.g., Fourier, Walsh, cosine/sine transform, etc.).

204   **Orthogonal functions:**
This subclass is indented under subclass 203. Subject matter wherein the function is orthogonal (transformations as applied to vector, matrix, linear and polynomial functions, for example).

205   **Frequency:**
This subclass is indented under subclass 201. Subject matter wherein the speech is represented by frequency.

206   **Specialized information:**
This subclass is indented under subclass 205. Subject matter wherein the frequency data is analyzed to identify specific speech information.

207   **Pitch:**
This subclass is indented under subclass 206. Subject matter wherein the specific speech information represents the predominant frequency of the speech.

208   **Voiced or unvoiced:**
This subclass is indented under subclass 207. Subject matter wherein the specific speech information represents the presence (voiced) or absence (unvoiced) of predominant frequency components.

209   **Formant:**
This subclass is indented under subclass 206. Subject matter wherein the specific speech information represents the frequency values of any of several resonance bands which determine the phonetic quality of a vowel sound.

210   **Silence decision:**
This subclass is indented under subclass 206. Subject matter wherein the specific speech information represent the presence or absence of speech.

211   **Time:**
This subclass is indented under subclass 201. Subject matter wherein the speech signal is represented using time (e.g., time measurements and energy measured over time).

212   **Pulse code modulation (PCM):**
This subclass is indented under subclass 211. Subject matter wherein the signal is sampled over time, and the magnitude of each sample is quantized and converted into a digital signal.

213   **Zero crossing:**
This subclass is indented under subclass 211. Subject matter wherein the zero crossings of the signal are used to measure time or frequency.

214   **Voiced or unvoiced:**
This subclass is indented under subclass 211. Subject matter wherein time measurements are used to determine the presence (voiced) or absence (unvoiced) of predominant frequency components.

215   **Silence decision:**
This subclass is indented under subclass 211. Subject matter wherein time measurements are used to determine the presence or absence of speech (e.g., pauses between words, etc.).

216   **Correlation function:**
This subclass is indented under subclass 211. Subject matter wherein analysis of speech is performed using relationships between time series samples.

217   **Autocorrelation:**
This subclass is indented under subclass 216. Subject matter wherein the relationships are between different speech samples taken from the same time series.

218   **Cross-correlation:**
This subclass is indented under subclass 216. Subject matter wherein the relationships are between speech samples taken from different time series.

219   **Linear prediction:**
This subclass is indented under subclass 201. Subject matter wherein input samples of speech are estimated from past samples of an input sequence.

220   **Analysis by synthesis:**
This subclass is indented under subclass 201. Subject matter wherein the speech signal is coded and corrected by the difference of the

decoded coded signal from the original <u>speech</u> signal.

221   Pattern matching vocoders:
This subclass is indented under subclass 201. Subject matter wherein <u>speech</u> signals are compared and matching patterns are encoded.

222   Vector quantization:
This subclass is indented under subclass 221. Subject matter wherein the encoding maps a sequence of continuous or discrete vectors into a digital sequence.

223   Excitation patterns:
This subclass is indented under subclass 221. Subject matter wherein the encoding models <u>speech</u> using representations including the primary frequency period or periods (e.g., <u>pitchexcitation</u>, multipulse <u>excitation</u>, etc.).

224   Normalizing:
This subclass is indented under subclass 201. Subject matter wherein modifications of the <u>speech</u> signal emphasize or deemphasize certain features (e.g., spectral slope, average power, etc.).

225   Gain control:
This subclass is indented under subclass 201. Subject matter wherein the <u>speech</u> is adjusted to maintain an average amplitude.

226   Noise:
This subclass is indented under subclass 201. Subject matter wherein the coding reduces the effects of undesired signal components.

227   Pre-transmission:
This subclass is indented under subclass 226. Subject matter wherein the coding precedes transmission.

228   Post-transmission:
This subclass is indented under subclass 226. Subject matter wherein decoding after transmission minimizes the effects of <u>noise</u> in the transmission path.

229   Adaptive bit allocation:
This subclass is indented under subclass 201. Subject matter wherein limited storage or transmission resources are allocated by giving more resources to areas containing more data and giving fewer resources to areas containing less data.

230   Quantization:
Subject matter under 201 wherein coded information is mapped into digital words described by binary symbols.

231   Recognition:
This subclass is indented under subclass 200. Subject matter wherein <u>speech</u> is separated into discrete components which are distinguished from one another.

232   Neural networks:
This subclass is indented under subclass 231. Subject matter using parallel distributed processing elements constructed in hardware or simulated in software.

233   Detect speech in noise:
This subclass is indented under subclass 231. Subject matter wherein the discrete components are distinguished from <u>noise</u>.

234   Normalizing:
This subclass is indented under subclass 231. Subject matter wherein the discrete components are modified to emphasize or deemphasize certain features (e.g., spectral slope, average power, etc.).

235   Speech to image:
This subclass is indented under subclass 231. Subject matter wherein the distinguished discrete components are converted into image output (e.g., text).

236   Specialized equations or comparisons:
This subclass is indented under subclass 231. Subject matter wherein the discrete components are distinguished using specific mathematical functions.

237   Correlation:
This subclass is indented under subclass 236. Subject matter wherein the specific function measures a <u>correlation</u> between discrete components (e.g., absolute magnitude difference functions (AMDF), <u>autocorrelation</u>, <u>cross-correlation</u>, etc.).

238 Distance:
This subclass is indented under subclass 236. Subject matter wherein the specific function measures the difference between discrete components.

239 Similarity:
This subclass is indented under subclass 236. Subject matter wherein the specific function measures the similarity between discrete components.

240 Probability:
This subclass is indented under subclass 236. Subject matter wherein the specific function uses probability to determine the occurrence of a discrete component.

241 Dynamic time warping:
This subclass is indented under subclass 236. Subject matter wherein time components of the discrete components are aligned with reference components (e.g., using dynamic programming).

242 Viterbi Trellis:
This subclass is indented under subclass 236. Subject matter wherein discrete components are distinguished by traversing possible paths through a time series.

243 Creating patterns for matching:
This subclass is indented under subclass 231. Subject matter including specific methods for registering the discrete components to be used as references.

244 Update patterns:
This subclass is indented under subclass 243. Subject matter wherein the references are modified to improve recognition (e.g., learning).

245 Clustering:
This subclass is indented under subclass 243. Subject matter wherein similar references are placed or divided into groups (e.g., K-means algorithm, nearest neighbor, etc.).

246 Voice recognition:
This subclass is indented under subclass 231. Subject matter wherein different voices are distinguished (e.g., speaker identification or verification).

247 Preliminary matching:
This subclass is indented under subclass 246. Subject matter using an initial comparison followed by a more detailed recognition.

248 Endpoint detection:
This subclass is indented under subclass 246. Subject matter including the identification of the beginning and ending points of speech sound segments.

249 Subportions:
This subclass is indented under subclass 246. Subject matter including separating speech into sound segments (e.g., utterances, words, phonemes, allophones, etc.).

250 Specialized models:
This subclass is indented under subclass 246. Subject matter including models which describe the interconnections between speech sound segments.

251 Word recognition:
This subclass is indented under subclass 231. Subject matter wherein different words are distinguished (i.e., the meaning of what is spoken).

252 Preliminary matching:
This subclass is indented under subclass 251. Subject matter using an initial comparison followed by a more detailed recognition.

253 Endpoint detection:
This subclass is indented under subclass 251. Subject matter identifying the beginning and ending points of words.

254 Subportions:
This subclass is indented under subclass 251. Subject matter identifying speech sound segments (e.g., phonemes, allophones, etc.).

255 Specialized models:
This subclass is indented under subclass 251. Subject matter including models which describe the interconnections between words or subportions of words.

256   Markov:
This subclass is indented under subclass 255. Subject matter wherein the models include states which represent <u>speech</u> sound portions and transitions which represent connections between <u>speech</u> sound portions (e.g., hidden Markov models, heuristic Markov models, etc.).

256.1   Hidden Markov Model (HMM) (EPO):
This subclass is indented under subclass 256. Subject matter wherein a Markov chain used in the recognition process has un-observable (hidden) states.

(1)   Note. The subject matter in this subclass is substantially the same in scope as ECLA (G10L 15/14M).

(2)   Note. The observation model itself is part of the stochastic process (Markov Chain) with an underlying stochastic process that is not directly observable, but can be observed through a set of stochastic processes that produce the sequence of observations.

(3)   Note. The HMM has different elements, including the following – number of states, the number of distinct observations per state, state transition probability distribution, the observation symbol probability distribution, and the initial state distribution.

(4)   Note. The manipulation of HMM s can be use in improving the probability of observation sequences, optimizing state sequences, or maximizing the probability of the state sequences.

(5)   Note. Subcategories to the types of HMM s include finite state, discrete versus continuous, mixture densities, autoregressive, null transition, tied states, and state duration.

256.2   Training of HMM (EPO):
This subclass is indented under subclass 256.1. Subject matter wherein the models include a learning process for recognizing speech data, e.g., the construction of a library of models for the words in a vocabulary, including the states.

(1)   Note. The subject matter in this subclass is substantially the same in scope as ECLA (G10L 15/14M1).

256.3   With insufficient amount of training data, e.g., state sharing, tying, and deleted interpolation (EPO):
This subclass is indented under subclass 256.2. Subject matter wherein intrinsic parameters of the HMM are modified to overcome lack of training data, and to simplify the model, e.g., state sharing, tying, and deleted interpolation.

(1)   Note. The subject matter in this subclass is substantially the same in scope as ECLA (G10L 15/14M1S).

(2)   Note. State sharing involves combining two or more separately trained models, one of which is more reliably trained than the other. The scenario in which this can happen is the case when we use tied states which forces "different" states to share an identical statistical characterization, effectively reducing the number of parameters in the model.

(3)   Note. Parameter tying involves setting up an equivalence relation between HMM parameters in different states. In this manner the number of independent parameters in the model is reduced and the parameter estimation becomes somewhat simpler and in some cases more reliable. Parameter tying is used when the observation density, for example, is known to be the same in two or more states.

(4)   Note. Deleted interpolation is a parameter method aimed to improve model reliability. The concept involves combining two or more separately trained models, one of which is more reliably trained than the other. The scenario in which this can happen is the case when we use tied states which forces "different" states to share an identical statistical characterization, effectively reducing the number of parameters in the model. The technique of deleted interpolation has been successfully applied to a number of problems in speech recognition, including the

estimation of trigram word probabilities for language models, and the estimation of HMM output probabilities for trigram phone models.

256.4 Duration modeling in HMM, e.g., semi HMM, segmental models, transition probabilities (EPO):
This subclass is indented under subclass 256.1. Subject matter wherein the HMM includes a duration state model for speech recognition, e.g., semi HMM's segmental models, and transition probabilities.

(1) Note. The subject matter in this subclass is substantially the same in scope as ECLA (G10L 15/14M2).

(2) Note. A semi- Markov HMM is like an HMM except each state can emit a sequence of observations.

(3) Note. Within a state segment models introduce dependency between frames via their common dependence on a trajectory. There may be only a single trajectory or a continuous mixture of trajectories. The probability distribution over the sequence of frames for a state, given the duration and trajectory, is then typically modeled as independent Gaussian distributions for each time step, centered on the trajectory.

(4) Note. Symbol emission probabilities are associated to the states and transition probabilities to the connections between them.

256.5 Hidden Markov (HM) Network (EPO):
This subclass is indented under subclass 256.1. Subject matter including a HMM structure wherein subgroups of HMM types are used to perform speech recognition.

(1) Note. The subject matter in this subclass is substantially the same in scope as ECLA (G10L 15/14M3).

(2) Note. Each subgroup can vary by type of model, model size, and observation symbols.

256.6 State Emission Probability (EPO):
This subclass is indented under subclass 256.1. Subject matter wherein the HMM contains probability density function such that an emission probability is calculated for each state within the model.

(1) Note. The subject matter in this subclass is substantially the same in scope as ECLA (G10L 15/14M4).

(2) Note. For each state j, and for each possible output, a probability that a particular output symbol o is observed in that state. This is represented by the function $b_j\ (o)$, which gives the probability that o is emitted in state j. This is called the emission probability.

256.7 Continuous density, e.g., Gaussian distribution, Laplace (EPO):
This subclass is indented under subclass 256.6. Subject matter wherein the HMM contains continuous probability density observation models for the purpose of avoiding possible signal degradation inherent with discrete representations of signals.

(1) Note. The subject matter in this subclass is substantially the same in scope as ECLA (G10L 15/14M4C).

256.8 Discrete density, e.g., Vector Quantization preprocessor, look up tables (EPO):
This subclass is indented under subclass 256.6. Subject matter wherein the HMM contains discrete probability density observation models which allows for the use of a discrete probability density within each state of the model.

(1) Note. The subject matter in this subclass is substantially the same in scope as ECLA (G10L 15/14M4D).

(2) Note. Discrete probability density is used when the state of the model is discrete (e.g. representing a letter of the alphabet). Vector quantization is used to model its state.

257   Natural language:
This subclass is indented under subclass 255. Subject matter wherein the models include grammatical constraints (e.g., syntax, etc.).

258   Synthesis:
This subclass is indented under subclass 200. Subject matter wherein component parts of a <u>speech</u> signal are combined to produce a synthetic <u>speech</u> output.

259   Neural networks:
This subclass is indented under subclass 258. Subject matter wherein synthetic <u>speech</u> output is formed using parallel distributed processing elements constructed in hardware or simulated in software.

260   Image to speech:
This subclass is indented under subclass 258. Subject matter wherein the component parts are related to image data (e.g., text to <u>speech</u>, etc.).

261   Vocal tract model:
This subclass is indented under subclass 258. Subject matter wherein the component parts model a human vocal tract.

262   Linear prediction:
This subclass is indented under subclass 258. Subject matter wherein the component parts are represented by coefficients derived from a sequence of past <u>speech</u> samples.

263   Correlation:
This subclass is indented under subclass 258. Subject matter wherein the component parts are represented by coefficients derived from relationships between time series <u>speech</u> samples.

264   Excitation:
This subclass is indented under subclass 258. Subject matter wherein the component parts are represented by the period of the primary frequency of the <u>speech</u> signal (e.g., <u>pitchexcitation</u>, multi-pulse <u>excitation</u>, etc.).

265   Interpolation:
This subclass is indented under subclass 258. Subject matter wherein the component parts are combined using estimates of intermediate values (e.g., waveform smoothing).

266   Specialized model:
This subclass is indented under subclass 258. Subject matter wherein the component parts are combined or linked together in a defined manner (e.g., Markov models, trees, tries (tables representing trees), graphs, etc.).

267   Time element:
This subclass is indented under subclass 258. Subject matter wherein the component parts comprise time based elements (e.g., words, phonemes, allophones, etc.).

268   Frequency element:
This subclass is indented under subclass 258. Subject matter wherein the component parts comprise frequency based elements (e.g., <u>pitch</u> variations, inflection, formants, etc.).

269   Transformation:
This subclass is indented under subclass 258. Subject matter wherein the component parts are restored to <u>speech</u> using specific mathematical functions (e.g., Fourier, Walsh, Hilbert, Z-transform, cosine/sine transforms, etc.).

270   Application:
This subclass is indented under subclass 200. Subject matter intended or designed for a specified use to which the <u>speech</u> signal processing is being applied.

270.1   Speech assisted network
This subclass is indented under subclass 270. Subject matter wherein a system that employs speech recognition or synthesis to control or to provide user feedback such that the processing of speech data may occur at various levels within a computer network.

(1)   Note.   Various levels of processing would include local or remote locations relative to the user in order to make use of available resources. For example, a local terminal might not have the necessary storage or processing power but this can be overcome by accessing resources

over a network. Such resources may include the raw processing power necessary for analysis and pattern matching as well as dictionaries having data relevant to large vocabularies and multiple languages.

(2)   Note. Nominal recitations of speech or audio in network applications are classified elsewhere.

SEE OR SEARCH CLASS:
348,   Television, subclasses 13 through 20 for 2-way interactive conferencing.
370,   Multiplex Communications, subclasses 229 through 240 for data flow congestion prevention or control, subclasses 260-269 for conferencing and subclass 351 for voice over internet.
375,   Pulse or Digital Communications, subclasses 354 and 356 for synchronizing data for streaming over the internet.
707,   Data Processing: Database, Data Mining, and File Management or Data Structures, subclasses 770, 966 through 974 and 999.010 for distributed databases searching and access.
709,   Electrical Computers and Digital processing systems: Multiple computer or Process Coordinating, subclasses 227 through 229 for network computer-to-computer connections.
715,   Data Processing: Presentation Processing of Document, Operator Interface Processing, and Screen Saver Display Processing, subclasses 234 through 242 for HTML, SGML documents.

271   **Handicap aid:**
This subclass is indented under subclass 270. Subject matter for assisting handicapped people (e.g., blind or speech impaired communication and control).

272   **Novelty item:**
This subclass is indented under subclass 270. Subject matter for novelty items (e.g., greeting cards, toys, etc.).

273   **Security system:**
This subclass is indented under subclass 270. Subject matter for providing security (e.g., limited access).

SEE OR SEARCH CLASS:
726,   Information Security, subclasses 1 through 36 for information security in computers or digital processing system.

274   **Warning/alarm system:**
This subclass is indented under subclass 270. Subject matter for providing an audible warning or alarm (e.g., multiple sensors, car gauges, etc.).

275   **Speech controlled system:**
This subclass is indented under subclass 270. Subject matter for controlling specific devices through speech or voice commands.

276   **Pattern display:**
This subclass is indented under subclass 270. Subject matter for providing visual output representing speech (e.g., computer displays of speech data).

277   **Translation:**
This subclass is indented under subclass 270. Subject matter for translating one language into another language.

278   **Sound editing:**
This subclass is indented under subclass 270. Subject matter wherein speech is edited using waveform portions or other representations of the sounds to be modified.

500   **AUDIO SIGNAL BANDWIDTH COMPRESSION OR EXPANSION:**
This subclass is indented under the class definition. Subject matter where there is either an expansion or reduction of the bandwidth required for transmission of a sound signal.

(1)   Note. This subclass and its indents provide for bandwidth compression or expansion of audio signals other than speech signals.

SEE OR SEARCH THIS CLASS, SUB-
CLASS:
200+,   for expansion or reduction of a <u>speech</u>
         signal's bandwidth.
503+,   for time compression or expansion of
         audio signals.

SEE OR SEARCH CLASS:
333,   Wave Transmission Lines and Net-
        works, subclass 14 for amplitude
        compression and expansion in a long
        transmission line.
348,   Television, subclasses 384.1 through
        440.1 for bandwidth reduction of an
        analog television signal.
358,   Facsimile and Static Presentation Pro-
        cessing, subclasses 426.01 through
        426.16 for bandwidth reduction of a
        facsimile signal.
360,   Dynamic Magnetic Information Stor-
        age or Retrieval, subclasses 8+ for the
        use of a magnetic recorder to alter the
        bandwidth of a signal.
369,   Dynamic Information Storage or
        Retrieval, subclass 60.01 for the use
        of a dynamic storage device to change
        the bandwidth of a signal.
370,   Multiplex Communications, subclass
        118 for bandwidth compression in a
        multiplex system.
375,   Pulse or Digital Communications,
        subclasses 240 through 241 for band-
        width compression or expansion of a
        pulse or digital signal, particularly
        subclasses 240.01-240.29 for digital
        television.
381,   Electrical Audio Signal Processing
        Systems and Devices, subclass 106
        for amplitude compression or expan-
        sion.
455,   Telecommunications, subclass 72 for
        message signal compression or expan-
        sion in an analog signal modulated
        carrier wave communication system.

501   With content reduction encoding:
This subclass is indented under subclass 500.
Subject matter combined with means to discard
and replace redundant information by a code
indicating what has been discarded.

SEE OR SEARCH CLASS:
341,   Coded Data Generation or Conver-
        sion, subclass 55 for content reduc-
        tion encoding, per se.

502   Delay line:
This subclass is indented under subclass 500.
Subject matter having means to cause a time
delay of a sound signal.

SEE OR SEARCH CLASS:
333,   Wave Transmission Lines and Net-
        works, subclasses 138 through 165
        for delay lines, per se.

**503   AUDIO SIGNAL TIME COMPRESSION
OR EXPANSION (E.G., RUN LENGTH
CODING):**
This subclass is indented under the class defini-
tion. Subject matter where there is either an
expansion or reduction of the time required for
transmission of a non<u>speech</u> sound signal.

SEE OR SEARCH THIS CLASS, SUB-
CLASS:
211+,   for expansion or reduction of the time
         required for transmission of a <u>speech</u>
         signal.
500+,   for frequency compression or expan-
         sion of a <u>nonspeech</u> audio signal.

SEE OR SEARCH CLASS:
358,   Facsimile and Static Presentation Pro-
        cessing, subclasses 426.01 through
        426.16 for time compression of a fac-
        simile signal.
360,   Dynamic Magnetic Information Stor-
        age or Retrieval, subclasses 8+ for the
        use of a magnetic recorder to alter the
        time duration of a recorded signal.
369,   Dynamic Information Storage or
        Retrieval, subclass 60.01 for the use
        of a dynamic storage device to alter
        the time duration of a recorded signal.
370,   Multiplex Communications, subclass
        109 for time compression or expan-
        sion in a time division multiplex sys-
        tem.
381,   Electrical Audio Signal Processing
        Systems and Devices, subclass 106
        for amplitude compression or expan-
        sion.

455, Telecommunications, subclass 72 for message signal compression or expansion in an analog signal modulated carrier wave communication system.

504 With content reduction encoding

This subclass is indented under subclass 503. Subject matter combined with means to discard and replace redundant information by a code indicating what has been discarded.

SEE OR SEARCH CLASS:

341, Coded Data Generation or Conversion, subclass 55 for content reduction encoding, per se.

E-SUBCLASSES

The E-subclasses in U.S. Class 704 provide for methods and devices for analyzing or synthesizing spoken language and for detecting, recognizing, or modifying speech signal characteristics.

E11.001 MISCELLANEOUS ANALYSIS OR DETECTION OF SPEECH CHARACTERISTICS (EPO):

This main group provides for processes and apparatus for analyzing or detecting speech characteristics not provided for elsewhere. This subclass is substantially the same in scope as ECLA classification G10L11/00.

E11.002 General speech analysis without concrete application (EPO):

This subclass is indented under subclass E11.001. This subclass is substantially the same in scope as ECLA classification G10L11/00A.

E11.003 Detection of presence or absence of speech signals (EPO):

This subclass is indented under subclass E11.001. This subclass is substantially the same in scope as ECLA classification G10L11/02.

E11.004 Voice/data decision (EPO):

This subclass is indented under subclass E11.003. This subclass is substantially the same in scope as ECLA classification G10L11/02D.

E11.005 End point detection (EPO):

This subclass is indented under subclass E11.003. This subclass is substantially the same in scope as ECLA classification G10L11/02E.

E11.006 Pitch determination of speech signals (EPO):

This subclass is indented under subclass E11.001. This subclass is substantially the same in scope as ECLA classification G10L11/04.

E11.007 Voiced-unvoiced decision (EPO):

This subclass is indented under subclass E11.001. This subclass is substantially the same in scope as ECLA classification G10L11/06.

E13.001 SPEECH SYNTHESIS; TEXT TO SPEECH SYSTEMS (EPO):

This main group provides for processes and apparatus for synthesizing speech. This subclass is substantially the same in scope as ECLA classification G10L13/00.

E13.002 Methods for producing synthetic speech; speech synthesizers (EPO):

This subclass is indented under subclass E13.001. This subclass is substantially the same in scope as ECLA classification G10L13/02.

E13.003 Concept-to-speech synthesizers; generation of natural phrases not from text but from machine-based concepts (EPO):

This subclass is indented under subclass E13.002. This subclass is substantially the same in scope as ECLA classification G10L13/02C.

E13.004 Sound editing, manipulating voice of the synthesizer (EPO):

This subclass is indented under subclass E13.002. This subclass is substantially the same in scope as ECLA classification G10L13/02E.

E13.005 Details of speech synthesis systems, e.g., synthesizer architecture, memory management, etc. (EPO):

This subclass is indented under subclass E13.001. This subclass is substantially the

same in scope as ECLA classification G10L13/ 04.

E13.006 Architecture of speech synthesizers (EPO):
This subclass is indented under subclass E13.005. This subclass is substantially the same in scope as ECLA classification G10L13/ 04A.

E13.007 Excitation (EPO):
This subclass is in dented under subclass E13.005. This subclass is substantially the same in scope as ECLA classification G10L13/ 04E.

E13.008 Systems using speech synthesizers (EPO):
This subclass is indented under subclass E13.005. This subclass is substantially the same in scope as ECLA classification G10L13/ 04U.

E13.009 Elementary speech units used in speech synthesizers; concatenation rules (EPO):
This subclass is indented under subclass E13.001. This subclass is substantially the same in scope as ECLA classification G10L13/ 06.

E13.01 Concatenation (EPO):
This subclass is indented under subclass E13.009. This subclass is substantially the same in scope as ECLA classification G10L13/ 06C.

E13.011 Text analysis, generation of parameters for speech synthesis out of text, e.g., grapheme to phoneme translation, prosody generation, stress, or intonation determination, etc. (EPO):
This subclass is indented under subclass E13.001. This subclass is substantially the same in scope as ECLA classification G10L13/ 08.

E13.012 Grapheme to phoneme, detection of language (EPO):
This subclass is indented under subclass E13.011. This subclass is substantially the same in scope as ECLA classification G10L13/ 08G.

E13.013 Prosody rules derived from text (EPO):
This subclass is indented under subclass E13.011. This subclass is substantially the

same in scope as ECLA classification G10L13/ 08P.

E13.014 Stress or intonation (EPO):
This subclass is indented under subclass E13.011. This subclass is substantially the same in scope as ECLA classification G10L13/ 08S.

E15.001 SPEECH RECOGNITION (EPO):
This main group provides for processes, systems, and apparatus for the recognition of speech, including training of speech recognition systems, language recognition, speech classification and search, speech-to-text systems, and evaluation or assessment of speech recognition systems. This subclass is substantially the same in scope as ECLA classification G10L15/00.

E15.002 Assessment or evaluation of speech recognition systems (EPO):
This subclass is indented under subclass E15.001. This subclass is substantially the same in scope as ECLA classification G10L15/ 00A.

E15.003 Language recognition (EPO):
This subclass is indented under subclass E15.001. This subclass is substantially the same in scope as ECLA classification G10L15/ 00L.

E15.004 Feature extraction for speech recognition; selection of recognition unit (EPO):
This subclass is indented under subclass E15.001. This subclass is substantially the same in scope as ECLA classification G10L15/ 02.

E15.005 Segmentation or word limit detection (EPO):
This subclass is indented under subclass E15.001. This subclass is substantially the same in scope as ECLA classification G10L15/ 04.

E15.006 Word boundary detection (EPO):
This subclass is indented under subclass E15.005. This subclass is substantially the same in scope as ECLA classification G10L15/ 04W.

E15.007 Creation of reference templates; training of speech recognition systems, e.g., adaptation to the characteristics of the speaker's voice, etc. (EPO):

This subclass is indented under subclass E15.001. This subclass is substantially the same in scope as ECLA classification G10L15/06.

E15.008 Training (EPO):

This subclass is indented under subclass E15.007. This subclass is substantially the same in scope as ECLA classification G10L15/06T.

E15.009 Adaptation (EPO):

This subclass is indented under subclass E15.007. This subclass is substantially the same in scope as ECLA classification G10L15/06A.

E15.01 In the frequency domain (EPO):

This subclass is indented under subclass E15.009. This subclass is substantially the same in scope as ECLA classification G10L15/06A1.

E15.011 To speaker (EPO):

This subclass is indented under subclass E15.009. This subclass is substantially the same in scope as ECLA classification G10L15/06A3.

E15.012 Supervised, i.e., under machine guidance (EPO):

This subclass is indented under subclass E15.011. This subclass is substantially the same in scope as ECLA classification G10L15/06A3S.

E15.013 Unsupervised (EPO):

This subclass is indented under subclass E15.011. This subclass is substantially the same in scope as ECLA classification G10L15/06A3U.

E15.014 Speech classification or search (EPO):

This subclass is indented under subclass E15.001. This subclass is substantially the same in scope as ECLA classification G10L15/08.

E15.015 Using distance or distortion measures between unknown speech and reference templates (EPO):

This subclass is indented under subclass E15.014. This subclass is substantially the same in scope as ECLA classification G10L15/10.

E15.016 Using dynamic programming techniques, e.g., Dynamic Time Warping (DTW), etc. (EPO):

This subclass is indented under subclass E15.014. This subclass is substantially the same in scope as ECLA classification G10L15/12.

E15.017 Using artificial neural networks (EPO):

This subclass is indented under subclass E15.014. This subclass is substantially the same in scope as ECLA classification G10L15/16.

E15.018 Using natural language modeling (EPO):

This subclass is indented under subclass E15.014. This subclass is substantially the same in scope as ECLA classification G10L15/18.

E15.019 Using context dependencies, e.g., language models, etc. (EPO):

This subclass is indented under subclass E15.018. This subclass is substantially the same in scope as ECLA classification G10L15/18C.

E15.02 Phonemic context, e.g., pronunciation rules, phonotactical constraints, phoneme n-grams, etc. (EPO):

This subclass is indented under subclass E15.019. This subclass is substantially the same in scope as ECLA classification G10L15/18C1.

E15.021 Grammatical context, e.g., disambiguation of the recognition hypotheses based on word sequence rules, etc. (EPO):

This subclass is indented under subclass E15.019. This subclass is substantially the same in scope as ECLA classification G10L15/18C2.

E15.022 Formal grammars, e.g., finite state automata, context free grammars, word networks, etc. (EPO):
This subclass is indented under subclass E15.021. This subclass is substantially the same in scope as ECLA classification G10L15/18C2F.

E15.023 Probabilistic grammars, e.g., word n-grams, etc. (EPO):
This subclass is indented under subclass E15.021. This subclass is substantially the same in scope as ECLA classification G10L15/18C2S.

E15.024 Semantic context, e.g., disambiguation of the recognition hypotheses based on word meaning, etc. (EPO):
This subclass is indented under subclass E15.019. This subclass is substantially the same in scope as ECLA classification G10L15/18C3.

E15.025 Using prosody or stress (EPO):
This subclass is indented under subclass E15.018. This subclass is substantially the same in scope as ECLA classification G10L15/18P.

E15.026 Parsing for meaning understanding (EPO):
This subclass is indented under subclass E15.018. This subclass is substantially the same in scope as ECLA classification G10L15/18U.

E15.027 Using statistical models, e.g., Hidden Marker Models (HMMs), etc. (EPO):
This subclass is indented under subclass E15.014. This subclass is substantially the same in scope as ECLA classification G10L15/14.

E15.028 Hidden Markov Models (HMMs) (EPO):
This subclass is indented under subclass E15.027. This subclass is substantially the same in scope as ECLA classification G10L15/14M.

E15.029 Training of Hidden Markov Models (HMMs) (EPO):
This subclass is indented under subclass E15.028. This subclass is substantially the same in scope as ECLA classification G10L15/14M1.

E15.03 With insufficient amount of training data, e.g., state sharing, tying, deleted interpolation, etc. (EPO):
This subclass is indented under subclass E15.029. This subclass is substantially the same in scope as ECLA classification G10L15/14M1S.

E15.031 Duration modeling in Hidden Markov Models (HMMs), e.g., semi- HMM, segmental models, transition probabilities, etc. (EPO):
This subclass is indented under subclass E15.028. This subclass is substantially the same in scope as ECLA classification G10L15/14M2.

E15.032 Hidden Markov Models (HMMs) network (EPO):
This subclass is indented under subclass E15.028. This subclass is substantially the same in scope as ECLA classification G10L15/14M3.

E15.033 State emission probabilities (EPO):
This subclass is indented under subclass E15.028. This subclass is substantially the same in scope as ECLA classification G10L15/14M4.

E15.034 Continuous densities, e.g., Gaussian distribution, Laplace, etc. (EPO):
This subclass is indented under subclass E15.033. This subclass is substantially the same in scope as ECLA classification G10L15/14M4C.

E15.035 Discrete densities, e.g., Vector Quantization preprocessor, look-up tables, etc. (EPO):
This subclass is indented under subclass E15.033. This subclass is substantially the same in scope as ECLA classification G10L15/14M4D.

E15.036 Neural Network (NN) as output probability estimator, e.g., hybrid HMM/NN, etc. (EPO):
This subclass is indented under subclass E15.033. This subclass is substantially the same in scope as ECLA classification G10L15/14M4N.

E15.037 Non-hidden Markov Model (EPO):
This subclass is indented under subclass E15.027. This subclass is substantially the same in scope as ECLA classification G10L15/14N.

E15.038 Recognition networks (EPO):
This subclass is indented under subclass E15.014. This subclass is substantially the same in scope as ECLA classification G10L15/08N.

E15.039 Speech recognition techniques for robustness in adverse environments, e.g., in noise, of stress induced speech, etc. (EPO):
This subclass is indented under subclass E15.001. This subclass is substantially the same in scope as ECLA classification G10L15/20.

SEE OR SEARCH THIS CLASS, SUBCLASS:
E21.002,for speech enhancement.

E15.04 Procedures used during a speech recognition process, e.g., man-machine dialogue, etc. (EPO):
This subclass is in dented under subclass E15.001. This subclass is substantially the same in scope as ECLA classification G10L15/22.

E15.041 Speech recognition using nonacoustical features, e.g., position of the lips, etc. (EPO):
This subclass is indented under subclass E15.001. This subclass is substantially the same in scope as ECLA classification G10L15/24.

E15.042 Using position of the lips, movement of the lips, or face analysis (EPO):
This subclass is indented under subclass E15.041. This subclass is substantially the same in scope as ECLA classification G10L15/24L.

E15.043 Speech to text systems (EPO):
This subclass is indented under subclass E15.001. This subclass is substantially the same in scope as ECLA classification G10L15/26.

E15.044 Speech recognition depending on application context, e.g., in a computer, etc. (EPO):
This subclass is in dented under subclass E15.043. This subclass is substantially the same in scope as ECLA classification G10L15/26C.

E15.045 Systems using speech recognizers (EPO):
This subclass is indented under subclass E15.043. This subclass is substantially the same in scope as ECLA classification G10L15/26A.

E15.046 Constructional details of speech recognition systems (EPO):
This subclass is indented under subclass E15.001. This subclass is substantially the same in scope as ECLA classification G10L15/28.

E15.047 Distributed recognition, e.g., in client-server systems for mobile phones or network applications, etc. (EPO):
This subclass is indented under subclass E15.046. This subclass is substantially the same in scope as ECLA classification G10L15/28D.

E15.048 Memory allocation or algorithm optimization to reduce hardware requirements (EPO):
This subclass is indented under subclass E15.046. This subclass is substantially the same in scope as ECLA classification G10L15/28H.

E15.049 Multiple recognizers used in sequence or in parallel; corresponding voting or score combination systems (EPO):
This subclass is indented under subclass E15.046. This subclass is substantially the same in scope as ECLA classification G10L15/28M.

E15.05 Recognizers for parallel processing (EPO):
This subclass is indented under subclass E15.046. This subclass is substantially the same in scope as ECLA classification G10L15/28P.

E17.001 SPEAKER IDENTIFICATION OR VERI-FICATION (EPO):
This main group provides for processes and apparatus for recognizing special voice characteristics, systems using speaker recognizers and details of speaker identification or verification processes or apparatus. This subclass is substantially the same in scope as ECLA classification G10L17/00.

E17.002 Recognition of special voice characteristics, e.g., for use in a lie detector; recognition of animal voices, etc. (EPO):
This subclass is indented under subclass E17.001. This subclass is substantially the same in scope as ECLA classification G10L17/00C.

E17.003 Systems using speaker recognizers (EPO):
This subclass is indented under subclass E17.001. This subclass is substantially the same in scope as ECLA classification G10L17/00U.

E17.004 Details (EPO):
This subclass is indented under subclass E17.001. This subclass is substantially the same in scope as ECLA classification G10L17/00B2.

E17.005 Preprocessing operations, e.g., segment selection, etc.; pattern representation or modeling, e.g. based on linear discriminant analysis (LDA), principal components, etc.; feature selection or extraction (EPO):
This subclass is indented under subclass E17.004. This subclass is substantially the same in scope as ECLA classification G10L17/00B2.

E17.006 Training, model building, enrollment (EPO):
This subclass is in dented under subclass E17.004. This subclass is substantially the same in scope as ECLA classification G10L17/00B6.

E17.007 Decision making techniques, pattern matching strategies (EPO):
This subclass is in dented under subclass E17.004. This subclass is substantially the same in scope as ECLA classification G10L17/00B8.

E17.008 Use of particular distance or distortion metric between probe pattern and reference templates (EPO):
This subclass is indented under subclass E17.007. This subclass is substantially the same in scope as ECLA classification G10L17/00B8D.

E17.009 Multimodal systems, i.e., based on the integration of multiple recognition engines or experts fusion (EPO):
This subclass is indented under subclass E17.007. This subclass is substantially the same in scope as ECLA classification G10L17/00B8M.

E17.01 Score normalization (EPO):
This subclass is indented under subclass E17.007. This subclass is substantially the same in scope as ECLA classification G10L17/00B8N.

E17.011 Use of phonemic categorization or speech recognition prior to speaker recognition or verification (EPO):
This subclass is indented under subclass E17.007. This subclass is substantially the same in scope as ECLA classification G10L17/00B8P.

E17.012 Hidden Markov Models (HMMs) (EPO):
This subclass is in dented under subclass E17.004. This subclass is substantially the same in scope as ECLA classification G10L17/00B14.

E17.013 Artificial neural networks, connectionist approaches (EPO):
This subclass is in dented under subclass E17.004. This subclass is substantially the same in scope as ECLA classification G10L17/00B16.

E17.014 Pattern transformations and operations aimed at increasing system robustness, e.g., against channel noise, different working conditions, etc. (EPO):
This subclass is in dented under subclass E17.004. This subclass is substantially the same in scope as ECLA classification G10L17/00B20.

E17.015 Interactive procedures, man-machine interface (EPO):

This subclass is in dented under subclass E17.004. This subclass is substantially the same in scope as ECLA classification G10L17/00B22.

E17.016 User prompted to utter a password or predefined text (EPO):

This subclass is indented under subclass E17.015. This subclass is substantially the same in scope as ECLA classification G10L17/00B22P.

E19.001 SPEECH OR AUDIO SIGNAL ANALYSIS-SYNTHESIS TECHNIQUES FOR REDUNDANCY REDUCTION, E.G., IN VOCODERS, ETC.; CODING OR DECODING OF SPEECH OR AUDIO SIGNALS; COMPRESSION OR EXPANSION OF SPEECH OR AUDIO SIGNALS, E.G., SOURCE-FILTER MODELS, PSYCHOACOUSTIC ANALYSIS, ETC. (EPO):

This main group provides for processes and apparatus for the coding, decoding, compression or expansion of speech or audio signals, including techniques for redundancy reduction, and psychoacoustic analysis. This subclass is substantially the same in scope as ECLA classification G10L19/00.

SEE OR SEARCH THIS CLASS, SUBCLASS:
E21.016, for time compression or expansion of speech waves.

E19.002 Perceptual measures for quality assessment (EPO):

This subclass is in dented under subclass E19.001. This subclass is substantially the same in scope as ECLA classification G10L19/00A.

E19.003 Correction of errors induced by the transmission channel, if related to the coding (EPO):

This subclass is in dented under subclass E19.001. This subclass is substantially the same in scope as ECLA classification G10L19/00E.

E19.004 Lossless audio signal coding; perfect reconstruction of coded audio signal by transmission of coding error (EPO):

This subclass is in dented under subclass E19.001. This subclass is substantially the same in scope as ECLA classification G10L19/00L.

E19.005 Multichannel audio signal coding and decoding, i.e., using interchannel correlation to reduce redundancies, e.g., joint-stereo, intensity-coding, matrixing, etc. (EPO):

This subclass is in dented under subclass E19.001. This subclass is substantially the same in scope as ECLA classification G10L19/00M.

E19.006 Comfort noise, silence coding (EPO):

This subclass is in dented under subclass E19.001. This subclass is substantially the same in scope as ECLA classification G10L19/00N.

E19.007 Speech coding using phonetic or linguistical decoding of the source; reconstruction using text-to-speech synthesis (EPO):

This subclass is in dented under subclass E19.001. This subclass is substantially the same in scope as ECLA classification G10L19/00S.

E19.008 Systems using vocoders (EPO):

This subclass is in dented under subclass E19.001. This subclass is substantially the same in scope as ECLA classification G10L19/00U.

E19.009 Audio watermarking, i.e., embedding inaudible data in the audio signal (EPO):

This subclass is in dented under subclass E19.001. This subclass is substantially the same in scope as ECLA classification G10L19/00W.

E19.01 Using spectral analysis, e.g., transform vocoders, subband vocoders, perceptual audio coders, psychoacoustically based lossy encoding, etc., e.g., MPEG audio, Dolby AC-3, etc. (EPO):

This subclass is indented under subclass E19.001. This subclass is substantially the same in scope as ECLA classification G10L19/02.

E19.011 Blocking, i.e., grouping of samples in time, choice of analysis window, overlap factor (EPO):
This subclass is in dented under subclass E19.01. This subclass is substantially the same in scope as ECLA classification G10L19/02B.

E19.012 Detection of transients and attacks for time/frequency resolution switching (EPO):
This subclass is in dented under subclass E19.011. This subclass is substantially the same in scope as ECLA classification G10L19/02B1.

E19.013 Noise substitution, i.e., substituting nontonal spectral components by noisy source (EPO):
This subclass is in dented under subclass E19.01. This subclass is substantially the same in scope as ECLA classification G10L19/02N.

SEE OR SEARCH THIS CLASS, SUBCLASS:
E19.006,for comfort noise for discontinuous speech transmission.

E19.014 Spectral prediction for pre-echo prevention; temporal noise shaping (TNS), e.g., in MPEG2 or MPEG4, etc. (EPO):
This subclass is in dented under subclass E19.01. This subclass is substantially the same in scope as ECLA classification G10L19/02P.

E19.015 Quantization or dequantization of spectral components (EPO):
This subclass is in dented under subclass E19.01. This subclass is substantially the same in scope as ECLA classification G10L19/02Q.

E19.016 Scalar quantization (EPO):
This subclass is in dented under subclass E19.015. This subclass is substantially the same in scope as ECLA classification G10L19/02Q2.

E19.017 Vector quantization, e.g., Twin-VQ audio, etc. (EPO):
This subclass is in dented under subclass E19.015. This subclass is substantially the same in scope as ECLA classification G10L19/02Q4.

E19.018 Using subband decomposition (EPO):
This subclass is indented under subclass E19.01. This subclass is substantially the same in scope as ECLA classification G10L19/02S.

E19.019 Subband vocoders (EPO):
This subclass is in dented under subclass E19.018. This subclass is substantially the same in scope as ECLA classification G10L19/02S1.

E19.02 Using orthogonal transformation (EPO):
This subclass is in dented under subclass E19.01. This subclass is substantially the same in scope as ECLA classification G10L19/02T.

E19.021 Using wavelet decomposition (EPO):
This subclass is in dented under subclass E19.02. This subclass is substantially the same in scope as ECLA classification G10L19/02T2.

E19.022 Dynamic bit allocation (EPO):
This subclass is in dented under subclass E19.001. This subclass is substantially the same in scope as ECLA classification G10L19/00B.

E19.023 Using predictive techniques; codecs based on source-filter modelization (EPO):
This subclass is in dented under subclass E19.001. This subclass is substantially the same in scope as ECLA classification G10L19/04.

E19.024 Determination or coding of the spectral characteristics, e.g., of the short-term prediction coefficients, etc. (EPO):
This subclass is in dented under subclass E19.023. This subclass is substantially the same in scope as ECLA classification G10L19/06.

E19.025 Line spectrum pair (LSP) vocoders (EPO):
This subclass is in dented under subclass E19.024. This subclass is substantially the same in scope as ECLA classification G10L19/06L.

E19.026 Determination or coding of the excitation function; determination or coding of the long-term prediction characteristics (EPO):
This subclass is in dented under subclass E19.023. This subclass is substantially the

same in scope as ECLA classification G10L19/08.

E19.027 Determination or coding of an excitation gain (EPO):
This subclass is in dented under subclass E19.026. This subclass is substantially the same in scope as ECLA classification G10L19/08G.

E19.028 Using mixed excitation model, e.g., MELP, MBE, Split band LPC, HVXC , etc. (EPO):
This subclass is in dented under subclass E19.026. This subclass is substantially the same in scope as ECLA classification G10L19/08M.

E19.029 Long-term prediction, i.e., removing periodical redundancies, e.g., adaptive codebook, pitch predictor, etc. (EPO):
This subclass is in dented under subclass E19.026. This subclass is substantially the same in scope as ECLA classification G10L19/08P.

E19.03 Using sinusoidal excitation model (EPO):
This subclass is in dented under subclass E19.026. This subclass is substantially the same in scope as ECLA classification G10L19/08S.

E19.031 Using prototype waveform decomposition or waveform interpolative coders (PWI) (EPO):
This subclass is in dented under subclass E19.026. This subclass is substantially the same in scope as ECLA classification G10L19/08W.

E19.032 Determination or coding of a multipulse excitation (EPO):
This subclass is in dented under subclass E19.026. This subclass is substantially the same in scope as ECLA classification G10L19/10.

E19.033 Algebraic codebook; sparse pulse excitation (EPO):
This subclass is in dented under subclass E19.032. This subclass is substantially the same in scope as ECLA classification G10L19/10A.

E19.034 Regular pulse excitation (EPO):
This subclass is indented under subclass E19.032. This subclass is substantially the same in scope as ECLA classification G10L19/10R.

E19.035 Determination or coding of a code excitation; code excited linear prediction (CELP) vocoders (EPO):
This subclass is in dented under subclass E19.026. This subclass is substantially the same in scope as ECLA classification G10L19/12.

E19.036 Pitch excitation, e.g., PSI-CELP (pitch synchronous innovation CELP), etc. (EPO):
This subclass is in dented under subclass E19.035. This subclass is substantially the same in scope as ECLA classification G10L19/12P.

E19.037 Residual excited linear prediction (RELP) (EPO):
This subclass is in dented under subclass E19.035. This subclass is substantially the same in scope as ECLA classification G10L19/12R.

E19.038 Vector sum excited linear prediction (VSELP) (EPO):
This subclass is in dented under subclass E19.035. This subclass is substantially the same in scope as ECLA classification G10L19/12V.

E19.039 Details of speech and audio coders (EPO):
This subclass is indented under subclass E19.023. This subclass is substantially the same in scope as ECLA classification G10L19/14.

E19.04 Vocoder architecture (EPO):
This subclass is in dented under subclass E19.039. This subclass is substantially the same in scope as ECLA classification G10L19/14A.

E19.041 Vocoders using multiple modes (EPO):
This subclass is in dented under subclass E19.04. This subclass is substantially the same in scope as ECLA classification G10L19/14A1.

E19.042 Using sound class specific coding, hybrid encoders, object- based coding (EPO):
This subclass is in dented under subclass E19.041. This subclass is substantially the same in scope as ECLA classification G10L19/14A1C.

E19.043 Mode decision, i.e., based on audio signal content versus external parameter (EPO):
This subclass is in dented under subclass E19.041. This subclass is substantially the same in scope as ECLA classification G10L19/14A1D.

E19.044 Variable rate or variable quality codecs, e.g., scalable representation encoding, etc. (EPO):
This subclass is in dented under subclass E19.041. This subclass is substantially the same in scope as ECLA classification G10L19/14A1R.

E19.045 Pre- or post-filtering (EPO):
This subclass is in dented under subclass E19.039. This subclass is substantially the same in scope as ECLA classification G10L19/14P.

E19.046 Pre-filtering, e.g., high frequency emphasis prior to encoding, etc. (EPO):
This subclass is in dented under subclass E19.045. This subclass is substantially the same in scope as ECLA classification G10L19/14P1.

E19.047 Post-filtering, e.g., pitch enhancement, formant emphasis for decoder, etc. (EPO):
This subclass is in dented under subclass E19.045. This subclass is substantially the same in scope as ECLA classification G10L19/14P2.

E19.048 Audio streaming, i.e., formatting and decoding of an encoded audio signal (EPO):
This subclass is in dented under subclass E19.039. This subclass is substantially the same in scope as ECLA classification G10L19/14S.

E19.049 Transcoding, i.e., converting between two coded representations avoiding cascaded coding-decoding (EPO):
This subclass is in dented under subclass E19.039. This subclass is substantially the same in scope as ECLA classification G10L19/14T.

E21.001 MODIFICATION OF AT LEAST ONE CHARACTERISTIC OF SPEECH WAVES (EPO):
This main group provides for processes and apparatus for modifying at least one characteristic of a speech signal. This subclass is substantially the same in scope as ECLA classification G10L21/00.

E21.002 Speech enhancement, e.g., noise reduction, echo cancellation, etc. (EPO):
This subclass is in dented under subclass E21.001. This subclass is substantially the same in scope as ECLA classification G10L21/02.

E21.003 Applications (EPO):
This subclass is in dented under subclass E21.002. This subclass is substantially the same in scope as ECLA classification G10L21/02A.

E21.004 Speech corrupted by noise (EPO):
This subclass is in dented under subclass E21.003. This subclass is substantially the same in scope as ECLA classification G10L21/02A1.

E21.005 Periodic noise (EPO):
This subclass is in dented under subclass E21.004. This subclass is substantially the same in scope as ECLA classification G10L21/02A1N.

E21.006 The noise being separate speech (EPO):
This subclass is indented under subclass E21.004. This subclass is substantially the same in scope as ECLA classification G10L21/02A1S.

E21.007 Speech corrupted by echo-reverberation (EPO):
This subclass is in dented under subclass E21.003. This subclass is substantially the

same in scope as ECLA classification G10L21/ 02A2.

E21.008 Speech corrupted by stress-Lombard effect (EPO):
This subclass is in dented under subclass E21.003. This subclass is substantially the same in scope as ECLA classification G10L21/ 02A3.

E21.009 Enhancement of intelligibility of clean or coded speech (EPO):
This subclass is in dented under subclass E21.003. This subclass is substantially the same in scope as ECLA classification G10L21/ 02A4.

E21.01 Enhancement of diverse speech (EPO):
This subclass is indented under subclass E21.009. This subclass is substantially the same in scope as ECLA classification G10L21/ 02A4D.

E21.011 Bandwidth extension taking place at the receiving side, e.g., generation of low- or high-frequency components, regeneration of spectral holes, etc. (EPO):
This subclass is indented under subclass E21.009. This subclass is substantially the same in scope as ECLA classification G10L21/ 02A4E.

E21.012 Separate reconstruction of interference and of speech signal (EPO):
This subclass is indented under subclass E21.003. This subclass is substantially the same in scope as ECLA classification G10L21/ 02A6.

E21.013 The interference being a separate speaker (EPO):
This subclass is in dented under subclass E21.012. This subclass is substantially the same in scope as ECLA classification G10L21/ 02A6S.

E21.014 Active noise canceling (EPO):
This subclass is in dented under subclass E21.003. This subclass is substantially the same in scope as ECLA classification G10L21/ 02A7.

E21.015 Public address system (EPO):
This subclass is in dented under subclass E21.003. This subclass is substantially the same in scope as ECLA classification G10L21/ 02A8.

E21.016 Suppression or repetition of time signal segments (EPO):
This subclass is indented under subclass E21.002. This subclass is substantially the same in scope as ECLA classification G10L21/ 02R.

E21.017 Time compression or expansion (EPO):
This subclass is in dented under subclass E21.001. This subclass is substantially the same in scope as ECLA classification G10L21/ 04.

E21.018 Suppression or repetition of time signal segments (EPO):
This subclass is in dented under subclass E21.017. This subclass is substantially the same in scope as ECLA classification G10L21/ 04R.

E21.019 Transformation of speech into a nonaudible representation, e.g., speech visualization, speech processing for tactile aids, etc. (EPO):
This subclass is in dented under subclass E21.001. This subclass is substantially the same in scope as ECLA classification G10L21/ 06.

E21.02 Synchronization of speech with image or synthesis of the lips movement from speech, e.g., for "talking heads," etc. (EPO):
This subclass is in dented under subclass E21.019. This subclass is substantially the same in scope as ECLA classification G10L21/ 06L.

END