CLAIMS

1. A system for real-time translation in communications, comprising

a plurality of terminals that include a first terminal and a second terminal;

at least one network, which is configured to allow media content to be exchanged or streamed between the terminals in real-time or near real-time, via a service route at a substantially synchronized or near-synchronized manner, wherein the media content includes original media content of at least one language that is inputted to each terminal, and translated media content of at least one language that is outputted at each terminal; and

at least one device that is either part of, or in communication with, the network, having at least one processor that operates a plurality of modules, with the device configured to receive the original media content from at least one terminal for generating the translated media content therefrom that is sent to at least one corresponding terminal via the service route;

wherein the original media content includes at least one original audio stream, at least one original video stream, or a combination thereof;

wherein the translated media content includes the original audio stream, the original video stream, at least one translated audio stream, at least one data file having text transcribed and/or translated from the original audio stream that is embeddable within the original video stream as one or more subtitle tracks, or a combination thereof, which are composited and synchronized by the device based on specifications of the service route of the network;

wherein the device receives a signal from the first terminal that is to be passed to the second terminal for receipt by the second terminal, during which the device determines the language of the translated media content for the first terminal, and as the device receives another signal from the second terminal that is to be passed to the first terminal that indicates call receipt by the second terminal, the device determines the language of the translated media content for the second terminal, in which communications are established between the terminals thereafter where the device then receives the original media content from the terminals for translation into translated media content based on the determined languages;

wherein the device is further configured to allow a change of both the language of the original media content and the language of the translated media content translated thereinto for one terminal , upon receipt of a message from the aforementioned terminal that relates to a change in its languages, during exchange or streaming of the media content

between the terminals in real-time or near real-time, in which the languages for the original media content and the translated media content are both different from previous; and

wherein the translated text is selectively corrected by the device in an automated manner, based on complete sentence segments, as the translated media content is exchanged or streamed between the terminals.

2. The system according to claim 1, wherein the network has a communication framework that is based on the Internet Protocol Multimedia Core Network Subsystem (IMS) architecture, wherein the framework is either a Voice over Long-Term Evolution (VoLTE) channel, a Video over Long-Term Evolution (ViLTE) channel, or a combination thereof.

3. The system according to claim 1 or 2, wherein the modules include a business logic module that is configured to either or both

authenticate user identity information of at least one corresponding user of one corresponding terminal; and

determine service subscription information of at least one corresponding user of one corresponding terminal that relates to any one or a combination of input language settings of the terminals, output language settings of the terminals, and an output format of the translated media content that are being exchanged or streamed in real-time or near real-time, at the terminals.

4. The system according to claim 3, wherein the output format of the translated media content include any one of

a first type of translated media content, which is composited and synchronized media stream from one terminal having the original audio stream, and a subtitled video stream having the original video stream and any one or a combination of one subtitle track transcribed from the original audio stream and translated subtitle tracks transcribed and translated from the original audio stream; or

a second type of translated media content, which is composited and synchronized media stream from the one terminal having the translated audio stream, and the subtitled video stream having the original video stream and any one or a combination of one subtitle track transcribed from the original audio stream and the translated subtitle tracks transcribed and translated from the original audio stream.

wherein the translated subtitle track, the translated audio stream, or a combination thereof, are of languages according to the output language settings of the terminal that receives the translated media content.

5. The system according to any one of the preceding claims, wherein the modules include an automated speech recognition module that is configured to

recognize an original language from the original audio stream; and

transcribe a first set of text in the original language from the original audio stream; and

generate at least one audio transcription data file having the transcribed first set of text, which is embeddable within the original video stream.

6. The system according to claim 5, wherein the modules include a comprehension and translation module that is configured to

receive audio transcription file having the transcribed first set of text;

translate the transcribed first set of text into a second set of text that is of at least one translated language; and

generate at least one translated transcription data file having the second set of text, which is embeddable within the original video stream.

7. The system according to claim 6, wherein the modules include a text-to-speech module that is configured to generate the translated audio stream based on the second set of text of the translated transcription data file.

8. The system according to claim 7, wherein the text-to-speech module is further configured identify and/or sample timbre(s) within the original audio stream to generate the translated audio stream that mimics the voice of one user at one terminal in which the original audio stream is transmitted therefrom.

9. The system according to any one of the preceding claims, wherein the modules include a service control module that is configured to determine receipt of at least one media translation service request message from at least one terminal within a pre-set time period, for the service control module, or in conjunction with the other modules, to either

initiate or continue exchange or streaming of translated media content between the terminals in real-time or near real-time; or

stop the exchange or streaming of translated media content between the terminals in real-time or near real-time, and optionally, delete the translated media content that is stored within the device.

10. The system according to claim 9, wherein the service control module is further configured to

delete the translated media content that is stored within the device.

11. The system according to any one of the preceding claims, wherein the modules include a signalling module and a media processing module which are configured to enable the original media content and the translated media content to be exchanged between corresponding terminals via the service route over the network in a synchronized manner in real-time or near real-time.

12. The system according to any one of the preceding claims, further comprising a cache unit configured to store any one or a combination of the original media stream, the translated media stream, and their associated media stream resources, for a pre-determined time period, as they are being exchanged or streamed between the terminals in real-time or near real-time, via the service route.

13. A method for real-time translation in communications, comprising the steps of

allowing media content to be exchanged or streamed between a plurality of terminals, which include a first terminal and a second terminal, in real-time or near real-time, via a service route at a substantially synchronized or near-synchronized manner, by at least one network, wherein the media content includes original media content of at least one language that is inputted to each terminal, and translated media content of at least one language that is outputted at each terminal, which involves

receiving the original media content from at least one terminal, by at least one device that is either part of, or in communication with, the network, wherein the device has at least one processor that operates a plurality of modules;

generating the translated media content from the original media content, by the device; and

sending the translated media content to at least one corresponding terminal via the service route, by the device;

wherein the original media content includes at least one original audio stream, at least one original video stream, or a combination thereof;

wherein the translated media content includes the original audio stream, the original video stream, at least one translated audio stream, at least one data file having text transcribed and/or translated from the original audio stream that is embeddable within the original video stream as one or more subtitle tracks, or a combination thereof, which are composited and synchronized by the device based on specifications of the service route of the network;

wherein the device receives a signal from the first terminal that is to be passed to the second terminal for receipt by the second terminal, during which the device determines the language of the translated media content for the first terminal, and as the device receives another signal from the second terminal that is to be passed to the first terminal that indicates call receipt by the second terminal, the device determines the language of the translated media content for the second terminal, in which communications are established between the terminals thereafter where the device then receives the original media content from the terminals for translation into translated media content based on the determined languages;

wherein the device is further configured to allow a change of both the language of the original media content and the language of the translated media content translated thereinto for one terminal , upon receipt of a message from the aforementioned terminal that relates to a change in its languages, , during exchange or streaming of the media content between the terminals in real-time or near real-time, in which the languages for the original media content and the translated media content are both different from previous; and

wherein the translated text is selectively corrected by the device in an automated manner, based on complete sentence segments, as the translated media content is exchanged or streamed between the terminals.

14. The method according to claim 13, wherein the network has a communication framework that is based on the Internet Protocol Multimedia Core Network Subsystem (IMS)

architecture, wherein the framework is either a Voice over Long-Term Evolution (VoLTE) channel, a Video over Long-Term Evolution (ViLTE) channel, or a combination thereof.

15. The method according to claim 13 or 14, further comprising the steps of

authenticating user identity information of at least one corresponding user of one corresponding terminal, by a business logic module;

determining service subscription information of at least one corresponding user of one corresponding terminal that relates to any one or a combination of input language settings of the terminals, output language settings of the terminals, and an output format of the translated media content that are being exchanged or streamed in real-time or near real-time at the terminals, by the business logic module;

or a combination thereof.

16. The method according to any one of claims 13 to 15, further comprising steps carried out by an automated speech recognition module, which include

recognising an original language from the original audio stream;

transcribing a first set of text in the original language from the original audio stream; and

generating at least one audio transcription data file having the transcribed first set of text, which is embeddable within the original video stream.

17. The method according to claim 16, further comprising steps carried out by a comprehension and translation module, which include

translating the transcribed first set of text into a second set of text that is of at least one translated language; and

generating at least one translated transcription data file having the second set of text, which is embeddable within the original video stream.

18. The method according to claim 17, further comprising the step of generate the translated audio stream based on the second set of text of the translated transcription data file by a text-to-speech module.

19. The method according to any one of claims 13 to 18, further comprising steps carried out by a service control module, which include

determining receipt of at least one media translation service request message from at least one terminal within a pre-set time period, for initiating or continuing exchange or streaming of translated media content between the terminals in real-time or near real-time;

determining absence of at least one media translation service request message from at least one terminal within a pre-set time period, for stopping the exchange or streaming of translated media content between the terminals in real-time or near real-time; and

deleting the translated media content that is stored within the device.

20. A computer-readable medium for real-time translation in communications, being part of a device that is either part of, or in communication with, a network that is configured to allow media content to be exchanged or streamed between a plurality of terminals, include a first terminal and a second terminal, in real-time or near real-time, via a service route, wherein the media content includes original media content of at least one language that is inputted to each terminal, and translated media content of at least one language that is outputted at each terminal, wherein the computer-readable medium stores modules including their instructions that, when executed by a processor of the device, causes the processor to

enable the device to receive the original media content from at least one terminal;

generate translated media content from the original media content; and

send the translated media content to corresponding terminals in real-time or near real-time, via the service route;

wherein the original media content includes at least one original audio stream, at least one original video stream, or a combination thereof; and

wherein the translated media content includes the original audio stream, the original video stream, at least one translated audio stream, at least one data file having text transcribed and/or translated from the original audio stream that is embeddable within the original video stream as one or more subtitle tracks, or a combination thereof, which are composited and synchronized by the device based on specifications of the service route of the network;

wherein the processor of the device, as instructed by the computer-readable medium, receives a signal from the first terminal that is to be passed to the second terminal for receipt by the second terminal, during which the device determines the language of the translated

media content for the first terminal, and as the device receives another signal from the second terminal that is to be passed to the first terminal that indicates call receipt by the second terminal, the device determines the language of the translated media content for the second terminal, in which communications are established between the terminals thereafter where the device then receives the original media content from the terminals for translation into translated media content based on the determined languages;

wherein processor of the device, as instructed by the computer-readable medium, is further configured to allow a change of both the language of the original media content and the language of the translated media content translated thereinto for one terminal , upon receipt of a message from the aforementioned terminal that relates to a change for its languages, during exchange or streaming of the media content between the terminals in real-time or near real-time, in which the languages for the original media content and the translated media content are both different from previous; and

wherein the translated text is selectively corrected by the processor of the device based on instructions from the computer-readable medium, in an automated manner, based on complete sentence segments, as the translated media content is exchanged or streamed between the terminals.