Claims:

1. A method of assisting in improving speech of a user in a designated language, the method including:

receiving text in the designated language, or in another language, from the user to be spoken by the user in the designated language;

processing the text to derive one or more expected properties of speech of the text in the designated language;

receiving audio of the user speaking the text in the designated language;

processing the audio to derive one or more properties of the user speaking the text;

comparing said one or more properties of the user speaking the text to corresponding ones of said one or more expected properties of said speech of the text to determine an accuracy of the user speaking the text; and

outputting feedback to the user indicative of the accuracy of the user speaking the text, wherein the method further includes:

parsing the text into one or more sentences or segments of words to be spoken by the user;

prompting the user to speak one of the sentences or segments of words at a time;

receiving audio of the user speaking the sentences or segments of the text; and

outputting said feedback to the user indicative of the accuracy of the user speaking the sentences or segments of the text.

2. A method as claimed in claim 1, further including outputting a prompt for the user to speak a first one of the sentences or segments and receiving audio of the user speaking the first one of the sentences or segments of the text.

3. A method as claimed in claim 2, further including outputting said feedback to the user indicative of the accuracy of the user speaking the first one of the sentences or segments of the text.

4. A method as claimed in claim 3, further including outputting a further prompt for the user to speak a second one of the sentences or segments after outputting the feedback to the user indicative of the accuracy of the user speaking the first one of the sentences or segments of the text.

5. A method as claimed in claim 2, wherein the prompt includes text of the first one of the sentences or segments to be displayed on a display of a user device to the user.

6 A method as claimed in claim 1, further including generating a speech recognition grammar of the text.

7. A method as claimed in any one of claims 1 to 6, further including processing the text to derive a synthesised expected speech of the text in the designated language.

8. A method as claimed in claim 7, further including outputting the synthesised expected speech of the text to the user via a speaker of a user device.

9. A method as claimed in any one of claims 1 to 8, wherein the accuracy includes confidence values associated with the accuracy of the one or more properties of the user speaking the text corresponding to the one or more expected properties of speech of the text.

10. A method as claimed in claim 9, wherein the feedback includes colours indicative of different confidence values of the accuracy.

11. A method as claimed in any one of claims 1 to 10, further including processing the text to derive predetermined prohibited words in the text so that the prohibited words can be censored.

12. A method as claimed in any one of claims 1 to 11, wherein the one or more expected properties of speech of the text and the one or more properties of the user speaking the text include at least one of pronunciation, fluency, and prosodic features.

13. A method as claimed in any one of claims 1 to 12, further including processing the audio to derive said one or more properties of the user speaking the text using speech recognition algorithms.

14. A method as claimed in any one of claims 1 to 13, further including processing the text to derive said one or more expected properties of the text using speech synthesis algorithms.

15. A system for assisting in improving speech of a user in a designated language, the system including a processor having:

an input module configured to:

receive text in the designated language, or in another language, from the user to be spoken by the user in the designated language; and

receive audio of the user speaking the text in the designated language;

a processing module configured to:

process the text to derive one or more expected properties of speech of the text in the designated language;

process the audio to derive one or more properties of the user speaking the text; and

compare said one or more properties of the user speaking the text to corresponding ones of said one or more expected properties of said speech of the text to determine an accuracy of the user speaking the text; and

an output module configured to output feedback to the user indicative of the accuracy of the user speaking the text, wherein

the processing module is further configured to parse the text into one or more sentences or segments of words;

the output module is further configured to output prompts for the user to speak one of the sentences or segments of words at a time;

the input module is further configured to receive audio of the user speaking the one or more sentences or segments of the text; and

the output module is further configured to output said feedback to the user indicative of the accuracy of the user speaking the sentences or segments of the text.

16. A system as claimed in claim 15, wherein the output module is further configured to output a prompt for the user to speak a first one of the sentences or segments.

17. A system as claimed in claim 16, wherein the input module is further configured to receive audio of the user speaking the first one of the sentences or segments of the text.

18. A system as claimed in claim 17, wherein the output module is further configured to output said feedback to the user indicative of the accuracy of the user speaking the first one of the sentences or segments of the text.

19. A system as claimed in claim 18, wherein the output module is further configured to output a further prompt for the user to speak a second one of the sentences or segments after outputting the feedback to the user indicative of the accuracy of the user speaking the first one of the sentences or segments of the text.

20. A system as claimed in claims 16, wherein the prompt includes text of the first one of the sentences or segments to be displayed on a display of a user device to the user.

21. A system as claimed in claim 15, wherein the processing module is further configured to generate a speech recognition grammar of the text.

22. A system as claimed in any one of claims 15 to 21, wherein the processing module is further configured to process the text to derive a synthesised expected speech of the text in the designated language.

23. A system as claimed in claim 22, wherein the output module is further configured to output the synthesised expected speech of the text to the user via a speaker of a user device.

24 A system as claimed in any one of claims 15 to 23, wherein the accuracy includes confidence values associated with the accuracy of the one or more properties of the user speaking the text corresponding to the one or more expected properties of speech of the text.

25. A system as claimed in claim 24, wherein the feedback includes colours indicative of different confidence values of the accuracy.

26. A system as claimed in any one of claims 15 to 25, wherein the processing module is further configured to process the text to derive predetermined prohibited words in the text so that the prohibited words can be censored.

27. A system as claimed in any one of claims 15 to 26, wherein the system further includes a memory for recording the audio of the user speaking the text.

28. A system as claimed in claim 27, wherein the processing module is further configured to process instances of the audio recorded in the memory to derive the one or more expected properties of speech of the text.

29. A system as claimed in any one of claims 15 to 28, wherein the one or more expected properties of speech of the text and the one or more properties of the user speaking the text include at least one of pronunciation, fluency, and prosodic features.

30. A system for assisting in improving speech of a user in a designated language, the system including:

a display configured to display text in the designated language, or in another language;

a text input means configured to input said text in the designated language, or in said another language by the user to be spoken by the user in the designated language;

a microphone configured to input audio of the user speaking the text in the designated language; and

a processor having:

an input module configured to:

receive the text inputted in said designated language, or in said another language; and

receive the audio of the user speaking the text in the designated language;

a processing module configured to:

process the text to derive one or more expected properties of speech of the text in the designated language;

process the audio to derive one or more properties of the user speaking the text; and

compare said one or more properties of the user speaking the text to corresponding ones of said one or more expected properties of said speech of the text to determine an accuracy of the user speaking the text; and

an output module configured to output feedback to the user indicative of the accuracy of the user speaking the text, wherein

the processing module is further configured to parse the text into one or more sentences or segments of words;

the output module is further configured to output prompts for the user to speak one of the sentences or segments of words at a time;

the input module is further configured to receive audio of the user speaking the one or more sentences or segments of the text; and

the output module is further configured to output said feedback to the user indicative of the accuracy of the user speaking the sentences or segments of the text.

31. A system as claimed in claim 30, wherein the display is further configured to display the feedback to the user.

32. A system as claimed in claim 31, further including a server including the processor in data communication over a network with a user device including the display, the text input means and the microphone.

33. Computer programme code which when executed implements the method of any one of claims 1 to 14.

34. A tangible computer readable medium including the programme code of claim 33.

35. A data file including the programme code of claim 33.