CLAIMS

We claim:

1. A method implemented by one or more processors, the method comprising:

processing audio data using one or more acoustic models to monitor for an occurrence of an invocation phrase configured to invoke an automated assistant, wherein the audio data is based on detection of spoken input of a user at a client device that includes an automated assistant interface for interacting with the automated assistant;

detecting, based on processing the audio data using the one or more acoustic models, the occurrence of the invocation phrase in a portion of the audio data;

determining, based on processing of the audio data using the one or more acoustic models or other processing of the audio data using one or more other acoustic models, that the portion of the audio data that includes the invocation phrase corresponds to a user profile that is accessible to the automated assistant;

identifying a language assigned to the user profile;

selecting a speech recognition model for the language, wherein selecting the speech recognition model for the language is based on determining the portion of the audio data corresponds to the user profile, and is based on identifying the language as assigned to the user profile;

using the selected speech recognition model to process a subsequent portion of the audio data that follows the portion of the audio data; and

causing the automated assistant to provide responsive content that is determined based on the processing of the subsequent portion using the selected speech recognition model.

2. The method of claim 1, further comprising: identifying an additional language assigned to the user profile; wherein selecting the speech recognition model comprises selecting the speech recognition model in lieu of an additional speech recognition model for the additional language.

3. The method of claim 2, where selecting the speech recognition model in lieu of the additional speech recognition model for the additional language comprises: identifying one or more contextual parameters associated with the audio data; and selecting the speech recognition model based on the one or more contextual parameters being more strongly associated, in the user profile, with the language than with the additional language.

4. The method of claim 3, wherein the one or more contextual parameters comprise an identifier of the client device.

5. The method of claim 3, wherein the one or more contextual parameters comprise one or multiple of: a time of day, a day of the week, and a location of the client device.

6. The method of claim 1, wherein selecting the speech recognition model for the language comprises selecting the speech recognition model based on at least one probability assigned to the language in the user profile, wherein the at least one probability is based on previous interactions, with the automated assistant, that are associated with the user profile.

7. The method of claim 6, wherein the at least one probability is associated with one or more contextual parameters, and further comprising: identifying that the one or more contextual parameters are associated with the audio data; and using the at least one probability in the selecting based on the at least one probability being associated with the one or more contextual parameters identified as associated with the audio data.

8. The method of claim 1, wherein using the selected speech recognition model to process the subsequent portion of the audio data that follows the portion of audio data comprises using the selected speech recognition model in generating text, in the language, that corresponds to the subsequent portion; and wherein causing the automated assistant to provide responsive content that is determined based on the processing of the subsequent portion using the selected speech recognition model comprises: generating the responsive content based on the text, and causing the automated assistant interface to render output that is based on the responsive content.

9. The method of claim 1, wherein the automated assistant is configured to access multiple different user profiles that are: available at the client device, and associated with multiple different users of the client device.

10. The method of claim 9, wherein the multiple different user profiles each identify one or more corresponding languages and a corresponding language probability for each of the corresponding languages, the corresponding language probabilities each based on previous

interactions between a corresponding one of the multiple different users and the automated assistant.

11. The method of claim 1, wherein using the selected speech recognition model to process the subsequent portion of the audio data comprises using the selected speech recognition model to generate first text in the language, and further comprising: identifying an additional language assigned to the user profile; selecting an additional speech recognition model for the additional language, wherein selecting the additional speech recognition model for the additional language is based on determining the portion of the audio data corresponds to the user profile, and is based on identifying the additional language as assigned to the user profile; using the selected additional speech recognition model to process the subsequent portion of the audio data that follows the portion of the audio data, wherein using the selected additional speech recognition model to process the subsequent portion of the audio data comprises using the selected speech recognition model to generate second text in the additional language; selecting the first text in the language, in lieu of the second text in the additional language; wherein causing the automated assistant to provide responsive content that is determined based on the processing of the subsequent portion using the selected speech recognition model comprises: causing, based on selecting the first text in the language, the automated assistant to provide responsive content that is determined based on the first text in the language.

12. A non-transitory computer readable storage medium configured to store instructions that, when executed by one or more processors, cause the one or more processors to perform operations that include:

processing audio data using one or more acoustic models to monitor for an occurrence of an invocation phrase configured to invoke an automated assistant, wherein the audio data is based on detection of spoken input of a user at a client device that includes an automated assistant interface for interacting with the automated assistant;

detecting, based on processing the audio data using the one or more acoustic models, the occurrence of the invocation phrase in a portion of the audio data;

determining, based on processing of the audio data using the one or more acoustic models or other processing of the audio data using one or more other acoustic models, that the portion of the audio data that includes the invocation phrase corresponds to a user profile that is accessible to the automated assistant;

identifying a language assigned to the user profile;

selecting a speech recognition model for the language, wherein selecting the speech recognition model for the language is based on determining the portion of the audio data corresponds to the user profile, and is based on identifying the language as assigned to the user profile;

using the selected speech recognition model to process a subsequent portion of the audio data that follows the portion of the audio data; and

causing the automated assistant to provide responsive content that is determined based on the processing of the subsequent portion using the selected speech recognition model.

13. The non-transitory computer readable storage medium of claim 12, wherein the operations further include: identifying an additional language assigned to the user profile; wherein selecting the speech recognition model comprises selecting the speech recognition model in lieu of an additional speech recognition model for the additional language.

14. The non-transitory computer readable storage medium of claim 13, where selecting the speech recognition model in lieu of the additional speech recognition model for the additional language comprises: identifying one or more contextual parameters associated with the audio data; and selecting the speech recognition model based on the one or more contextual parameters being more strongly associated, in the user profile, with the language than with the additional language.

15. The non-transitory computer readable storage medium of claim 12, wherein using the selected speech recognition model to process the subsequent portion of the audio data that follows the portion of audio data comprises using the selected speech recognition model in generating text, in the language, that corresponds to the subsequent portion; and wherein causing the automated assistant to provide responsive content that is determined based on the processing of the subsequent portion using the selected speech recognition model comprises: generating the responsive content based on the text, and causing the automated assistant interface to render output that is based on the responsive content.

16. The non-transitory computer readable storage medium of claim 12, wherein the automated assistant is configured to access multiple different user profiles that are: available at the client device, and associated with multiple different users of the client device.

17. A system, comprising:

one or more processors; and

memory storing instructions that, when executed by the one or more processors, cause the one or more processors to perform operations that include:

processing audio data using one or more acoustic models to monitor for an occurrence of an invocation phrase configured to invoke an automated assistant, wherein the audio data is based on detection of spoken input of a user at a client device that includes an automated assistant interface for interacting with the automated assistant;

detecting, based on processing the audio data using the one or more acoustic models, the occurrence of the invocation phrase in a portion of the audio data;

determining, based on processing of the audio data using the one or more acoustic models or other processing of the audio data using one or more other acoustic models, that the portion of the audio data that includes the invocation phrase corresponds to a user profile that is accessible to the automated assistant;

identifying a language assigned to the user profile;

selecting a speech recognition model for the language, wherein selecting the speech recognition model for the language is based on determining the portion of the audio data corresponds to the user profile, and is based on identifying the language as assigned to the user profile;

using the selected speech recognition model to process a subsequent portion of the audio data that follows the portion of the audio data; and

causing the automated assistant to provide responsive content that is determined based on the processing of the subsequent portion using the selected speech recognition model.

18. The system of claim 17, wherein the operations further include: identifying an additional language assigned to the user profile; wherein selecting the speech recognition model comprises selecting the speech recognition model in lieu of an additional speech recognition model for the additional language.

19. The system of claim 18, where selecting the speech recognition model in lieu of the additional speech recognition model for the additional language comprises: identifying one or more contextual parameters associated with the audio data; and selecting the speech recognition model based on the one or more contextual parameters being more strongly associated, in the user profile, with the language than with the additional language.

20. The system of claim 17, wherein selecting the speech recognition model for the language comprises selecting the speech recognition model based on at least one probability assigned to the language in the user profile, wherein the at least one probability is based on previous interactions, with the automated assistant, that are associated with the user profile.