[0575] While the methods and systems have been described herein with respect to media content (e.g., music content, video content), the methods and systems described herein may be applied to a variety of content which may have associated audio that can be played by a media playback system. For example, pre-recorded sounds which might not be part of a music catalog may be played in response to a voice input. One example is the voice input "what does a nightingale sound like?". The networked microphone system's response to this voice input might not be music content with an identifier and may instead be a short audio clip. The media playback system may receive information associated with playing back the short audio clip (e.g., storage address, link, URL, file) and a media playback system command to play the short audio clip. Other examples are possible including podcasts, news clips, notification sounds, alarms, etc.

IV. Conclusion

[0576] The description above discloses, among other things, various example systems, methods, apparatus, and articles of manufacture including, among other components, firmware and/or software executed on hardware. It is understood that such examples are merely illustrative and should not be considered as limiting. For example, it is contemplated that any or all of the firmware, hardware, and/or software aspects or components can be embodied exclusively in hardware, exclusively in software, exclusively in firmware, or in any combination of hardware, software, and/or firmware. Accordingly, the examples provided are not the only way(s) to implement such systems, methods, apparatus, and/or articles of manufacture.

[0577] Additionally, references herein to "embodiment" means that a particular feature, structure, or characteristic described in connection with the embodiment can be included in at least one example embodiment of an invention. The appearances of this phrase in various places in the specification are not necessarily all referring to the same embodiment, nor are separate or alternative embodiments mutually exclusive of other embodiments. As such, the embodiments described herein, explicitly and implicitly understood by one skilled in the art, can be combined with other embodiments.

[0578] The specification is presented largely in terms of illustrative environments, systems, procedures, steps, logic blocks, processing, and other symbolic representations that directly or indirectly resemble the operations of data processing devices coupled to networks. These process descriptions and representations are typically used by those skilled in the art to most effectively convey the substance of their work to others skilled in the art. Numerous specific details are set forth to provide a thorough understanding of the present disclosure. However, it is understood to those skilled in the art that certain embodiments of the present disclosure can be practiced without certain, specific details. In other instances, well known methods, procedures, components, and circuitry have not been described in detail to avoid unnecessarily obscuring aspects of the embodiments. Accordingly, the scope of the present disclosure is defined by the appended claims rather than the forgoing description of embodiments.

[0579] When any of the appended claims are read to cover a purely software and/or firmware implementation, at least one of the elements in at least one example is hereby expressly defined to include a tangible, non-transitory medium such as a memory, DVD, CD, Blu-ray, and so on, storing the software and/or firmware.

1. A method comprising:

while one or more playback devices in a media playback system are playing audio content, receiving a voice command via a networked microphone device associated with the media playback system;

selecting one playback device of the one or more playback devices of the media playback system to play an audio response to the voice command based on (i) a direction of the voice command relative to the networked microphone device and (ii) a proximity of the voice command relative to the networked microphone device;

determining a directionality for the audio response to the voice command based on the direction of the voice command relative to the networked microphone device; and

while the selected one playback device is playing the audio content at a first volume level, causing the selected one playback device to (i) reduce its volume to a second volume level lower than the first volume level and (ii) while playing the audio content at the second volume level, play the audio response according to the determined directionality for the audio response.

2. The method of claim **1**, wherein selecting the selected one playback device comprises causing the networked microphone device to select a determined audio response zone, and wherein causing the selected one playback device to play the audio response comprises causing the selected one playback device to play the audio response.

3. The method of claim **1**, further comprising:

receiving, device from a controller device, an indication of an audio response zone to which the networked microphone device is assigned.

4. The method of claim **1**, further comprising:

sending, to a computing device, an indication of an audio response zone to which the networked microphone device is assigned.

5. The method of claim **1**, further comprising:

applying acoustics of an environment in which the networked microphone device is placed to the received voice command.

6. The method of claim **1**, further comprising:

receiving an indication of audio content;

determining at least one music service from a plurality of music services that supports the audio content; and

causing the at least one music service to transmit the audio content to the media playback system.

7. The method of claim **1**, further comprising:

establishing a metadata exchange channel between the media playback system and the networked microphone device.

8. The method of claim **1**, wherein the media playback system comprises a fallback device, the method further comprising:

before causing the selected one playback device to play the audio response, determining that another playback device is not available to play the audio response.

9. The method of claim **1**, further comprising:

determining that the received voice command was received from a registered user of the media playback system; and