WHAT IS CLAIMED IS:

1. A method performed by a voice-enabled computer system, the method comprising:

obtaining a current nonce;

operating a speaker of the voice-enabled computer system to produce a current acoustic signature based on the current nonce;

while producing the current acoustic signature, operating a microphone of the voice-enabled computer system to record audio that includes the current acoustic signature and a speech input; and

validating the recorded audio based at least in part on the current acoustic signature; and

in the event that the recorded audio is validated:

processing the recorded audio to extract the speech input;

processing the speech input to determine an action to perform; and

performing the action.

2. The method of claim 1, wherein validating the recorded audio includes:

determining whether the recorded audio includes the current acoustic signature; and

determining whether the recorded audio includes a different acoustic signature based on a previous nonce,

wherein the recorded audio is validated in the event that the recorded audio includes the current acoustic signature and does not include the different acoustic signature.

3. The method of claim 1 wherein operating the speaker of the voice-enabled computer system to produce the current acoustic signature based on the current nonce includes:

defining a frequency-hopping pattern for the current acoustic signature based on a first portion of the current nonce;

defining a modulation sequence based on a second portion of the current nonce; and

generating sound at the speaker based on the frequency-hopping pattern and the modulation sequence.

4. The method of claim 3 wherein the frequency-hopping pattern includes frequencies in a range that overlaps a vocal range of an authorized user.

5. The method of claim 1 further comprising:

prompting a user to speak such that the user speaks while the current acoustic signature is being produced.

6. The method of claim 1 wherein the speech input includes a pass phrase and processing the speech input includes performing voice-based identity verification on the speech input.

7. The method of claim 1 wherein processing the recorded audio includes using the current nonce to filter out the current acoustic signature from the recorded audio.

8. A server computer comprising:

a processor;

a memory; and

a computer readable medium coupled to the processor, the computer readable medium having stored therein code executable by the processor to implement a method comprising:

generating a nonce;

providing the nonce to a voice-enabled client having a speaker and a microphone and capable of generating an acoustic signature based on the nonce;

receiving, from the voice-enabled client, an audio recording;

validating the audio recording based at least in part on detecting the acoustic signature in the audio recording; and

processing user speech in the audio recording only if the audio recording is validated.

9. The server computer of claim 8 wherein the method further comprises adding the nonce to a list of previous nonces after validating the audio recording.

10. The server computer of claim 9 wherein validating the audio recording further includes:

determining, based on the list of previous nonces, whether an old acoustic signature is present in the audio recording,

wherein the audio recording is not validated if an old acoustic signature is present.

11. The server computer of claim 8 wherein processing the user speech in the audio recording includes filtering out the acoustic signature.

12. The server computer of claim 11 wherein processing the user speech in the audio recording includes performing a voice-based identity verification after filtering out the acoustic signature.

13. The server computer of claim 8 wherein providing the nonce to the voice-enabled client is performed in response to a request from the voice-enabled client that involves sensitive information.

14. A voice-enabled computer system comprising:

a processor;

a memory;

a speaker operable by the processor;

a microphone operable by the processor; and

a computer readable medium coupled to the processor, the computer readable medium having stored therein code executable by the processor to implement a method comprising:

obtaining a nonce;

operating the speaker to produce an acoustic signature based on the nonce; and

while producing the acoustic signature, operating the microphone to record an audio recording that includes the acoustic signature and a speech input.

15. The voice-enabled computer system of claim 14 wherein the method further comprises:

validating the audio recording based at least in part on detecting the acoustic signature; and

processing the speech input only in the event that the audio recording is validated.

16. The voice-enabled computer system of claim 14 wherein the nonce is obtained from a server computer and the method further comprises:

sending the audio recording to the server computer for processing;

receiving a response from the server computer; and

providing the response to a user.

17. The voice-enabled computer system of claim 14 wherein operating the speaker to produce the acoustic signature based on the nonce includes:

using the nonce to define a frequency-hopping pattern and a modulation sequence for the acoustic signature; and

generating sound at the speaker based on the frequency-hopping pattern and the modulation sequence.

18. The voice-enabled computer system of claim 14 wherein the acoustic signature includes frequencies in a range that overlaps a vocal range of a human voice.

19. The voice-enabled computer system of claim 14 wherein obtaining the nonce and operating the speaker to produce the acoustic signature are performed in response to receiving a user request that entails voice-based identity verification.

20. The voice-enabled computer system of claim 14 wherein the method further comprises:

detecting a speech input corresponding to an activation phrase for the voice-enabled computer system,

wherein obtaining the current nonce and operating the speaker are performed in response to detecting the speech input corresponding to the activation phrase.