Web Speech API

This is an experimental technology
Check the Browser compatibility table carefully before using this in production.

The Web Speech API enables you to incorporate voice data into web apps. The Web Speech API has two parts: SpeechSynthesis (Text-to-Speech), and SpeechRecognition (Asynchronous Speech Recognition.)

Web Speech Concepts and Usage

The Web Speech API makes web apps able to handle voice data. There are two components to this API:

- Speech recognition is accessed via the `SpeechRecognition` interface, which provides the ability to recognize voice context from an audio input (normally via the device's default speech recognition service) and respond appropriately. Generally you'll use the interface's constructor to create a new `SpeechRecognition` object, which has a number of event handlers available for detecting when speech is input through the device's microphone. The `SpeechGrammar` interface represents a container for a particular set of grammar that your app should recognise. Grammar is defined using JSpeech Grammar Format (**JSGF**.)
- Speech synthesis is accessed via the `SpeechSynthesis` interface, a text-to-speech component that allows programs to read out their text content (normally via the device's default speech synthesiser.) Different voice types are represented by `SpeechSynthesisVoice` objects, and different parts of text that you want to be spoken are represented by `SpeechSynthesisUtterance` objects. You can get these spoken by passing them to the `SpeechSynthesis.speak()` method.

For more details on using these features, see Using the Web Speech API.

Web Speech API Interfaces

Speech recognition

`SpeechRecognition`
: The controller interface for the recognition service; this also handles the `SpeechRecognitionEvent` sent from the recognition service.

`SpeechRecognitionAlternative`
: Represents a single word that has been recognised by the speech recognition service.

`SpeechRecognitionError`
: Represents error messages from the recognition service.

`SpeechRecognitionEvent`

The event object for the `result` and `nomatch` events, and contains all the data associated with an interim or final speech recognition result.

`SpeechGrammar`

The words or patterns of words that we want the recognition service to recognize.

`SpeechGrammarList`

Represents a list of `SpeechGrammar` objects.

`SpeechRecognitionResult`

Represents a single recognition match, which may contain multiple `SpeechRecognitionAlternative` objects.

`SpeechRecognitionResultList`

Represents a list of `SpeechRecognitionResult` objects, or a single one if results are being captured in `continuous` mode.

Speech synthesis

`SpeechSynthesis`

The controller interface for the speech service; this can be used to retrieve information about the synthesis voices available on the device, start and pause speech, and other commands besides.

`SpeechSynthesisErrorEvent`

Contains information about any errors that occur while processing `SpeechSynthesisUtterance` objects in the speech service.

`SpeechSynthesisEvent`

Contains information about the current state of `SpeechSynthesisUtterance` objects that have been processed in the speech service.

`SpeechSynthesisUtterance`

Represents a speech request. It contains the content the speech service should read and information about how to read it (e.g. language, pitch and volume.)

`SpeechSynthesisVoice`

Represents a voice that the system supports. Every `SpeechSynthesisVoice` has its own relative speech service including information about language, name and URI.

`Window.speechSynthesis`

Specced out as part of a `[NoInterfaceObject]` interface called `SpeechSynthesisGetter`, and Implemented by the `Window` object, the `speechSynthesis` property provides access to the `SpeechSynthesis` controller, and therefore the entry point to speech synthesis functionality.

Examples

The Web Speech API repo on GitHub contains demos to illustrate speech recognition and synthesis.

Specifications

Specification	Status	Comment
Web Speech API	D Draft	Initial definition

Browser compatibility

`SpeechRecognition`

Update compatibility data on GitHub

Chrome	Edge	Firefox	Internet Explorer	Opera	Safari	Android webview	Chrome for Android	Firefox for Android	Opera for Android	Safari on iOS	Samsung Internet
`SpeechRecognition`											
33 -x- *	≤79 -x- *	No	No	No	No	4.4.3 -x- *	33 -x- *	No	No	No	2.0 -x- *

What are we missing?

Full support — No support

Experimental. Expect behavior to change in the future. — See implementation notes.

-x- Requires a vendor prefix or different name for use.

`SpeechSynthesis`

Update compatibility data on GitHub

Chrome	Edge	Firefox	Internet Explorer	Opera	Safari	Android webview	Chrome for Android	Firefox for Android	Opera for Android	Safari on iOS	Samsung Internet
`SpeechSynthesis`											
33	≤18	49	No	21	7	4.4.3	33	62	No	7	3.0

Full support — No support

Experimental. Expect behavior to change in the future. — User must explicitly enable this feature.

See also

- Using the Web Speech API
- SitePoint article
- HTML5Rocks article
- Demo [aurelio.audero.it]

Last modified: Mar 18, 2019, by MDN contributors

Related Topics

- Interfaces
  - SpeechGrammar
  - SpeechGrammarList
  - SpeechRecognition
  - SpeechRecognitionAlternative
  - SpeechRecognitionError
  - SpeechRecognitionEvent
  - SpeechRecognitionResult
  - SpeechRecognitionResultList
  - SpeechSynthesis
  - SpeechSynthesisErrorEvent
  - SpeechSynthesisEvent
  - SpeechSynthesisUtterance
  - SpeechSynthesisVoice