WebSocket Message Specification

Each message exchanged over the WebSocket connection is sent and received as a **JSON string**.
 The following message types are supported:

- **Connected**

- **Start**

- **Media**

- **DTMF**

- **Stop**

- **Mark** *(only in bidirectional streaming)*

- **Clear**

Connected Message

Sent immediately after the WebSocket connection is successfully established.

```
{
  "event": "connected"
}
```

Start Message

Contains information about the stream parameters.
 Sent once, right after the `connected` message.

```json
{
  "event": "start",
  "sequence_number": 1,
  "stream_sid": "<stream sid>",
  "start": {
    "stream_sid": "<>",
    "call_sid": "",
    "media_format": {
      "encoding": "<>",
      "sample_rate": "<>",
      "bit_rate": "<>"
    }
  }
}
```

Media Message

Encapsulates audio packets transmitted in the stream.

```json
{
  "event": "media",
  "sequence_number": 3,
  "stream_sid": "<stream sid>",
  "media": {
    "chunk": 2,
    "timestamp": "10",
    "payload": "<>"
  }
}
```

Notes:

- `media.chunk`: Index of the current media chunk

- `media.timestamp`: Timestamp in milliseconds since the start of the stream

- `payload`: Base64-encoded raw audio (SLIN – 16-bit, 8kHz, mono PCM, little-endian)

DTMF Message

Sent when the user presses a key during the session.
 Supported only for **bidirectional streaming**.

```
{
  "event": "dtmf",
  "sequence_number": 1,
  "stream_sid": "<stream sid>",
  "dtmf": {
    "duration": "<duration in ms>",
    "digit": "<digit>"
  }
}
```

Stop Message

Indicates that the stream has been stopped or the call has ended.

```
{
  "event": "stop",
  "sequence_number": 10,
  "stream_sid": "<stream sid>",
  "stop": {
    "call_sid": "<>",
    "account_sid": "<>",
    "reason": "stopped or callended"
  }
}
```

Mark Message

Used only to track when media playback is complete.

Example (Received or Sent)

```
{
```

```
 "event": "mark",
 "sequence_number": 15,
 "stream_sid": "<stream sid>",
 "mark": {
   "name": "<label>"
 }
}
```

Clear Message

Used to clear audio data that was previously sent but not yet played.

```
{
  "event": "clear",
  "stream_sid": "<stream sid>"
}
```

Media Format

All audio payloads use the following format:

- **Encoding:** raw/slin (16-bit, 8kHz, mono PCM, little-endian)

- **Transmission:** Base64 encoded

- **Bidirectional Streams:** The same format must be used for sending audio back to the caller.

Event Structure Reference

Field Name	Type	JSON Key	Optional	Description

Event	string	`event`	No	Type of event (`start`, `media`, `stop`, `dtmf`, etc.)
StreamSID	string	`stream_sid`	Yes	Unique identifier for the stream session
SequenceNumber	string	`sequence_number`	Yes	Sequence number for message ordering
Start	object	`start`	Yes	Present when the event is a `start` event
Media	object	`media`	Yes	Present when the event is a `media` event
Stop	object	`stop`	Yes	Present when the event is a `stop` event
Mark	object	`mark`	Yes	Present when the event is a `mark` event
Dtmf	object	`dtmf`	Yes	Present when the event is a `dtmf` event