

Powered by Universal Speech Solutions LLC



AWS Lex Plugin

Usage Guide

Revision: 1

Created: October 15, 2018

Last updated: September 15, 2018

Author: Arsen Chaloyan

Table of Contents

1	Overview.....	4
1.1	Installation	4
1.2	Applicable Versions.....	4
2	Supported Features.....	5
2.1	MRCP Methods	5
2.2	MRCP Events	5
2.3	MRCP Header Fields	5
2.4	Grammars.....	6
2.5	Results.....	6
3	Configuration Format.....	7
3.1	Document.....	7
3.2	Streaming Recognition	8
3.3	Results.....	9
3.4	Utterance Manager.....	10
3.5	RDR Manager	11
3.6	Monitoring Agent	12
3.7	Usage Change Handler	13
3.8	Usage Refresh Handler	13
3.9	License Server.....	14
4	Configuration Steps	16
4.1	Using Default Configuration.....	16
4.2	Using with Polly	16
4.3	Specifying Lex Bot	16
4.4	Specifying Recognition Language.....	16
4.5	Specifying Sampling Rate.....	17
4.6	Specifying Speech Input Parameters	17
4.7	Specifying DTMF Input Parameters.....	17
4.8	Specifying No-Input and Recognition Timeouts	18
4.9	Maintaining Utterances.....	18
4.10	Maintaining Recognition Details Records	19
5	Recognition Grammars and Results.....	20
5.1	Using Built-in Speech Grammar.....	20
5.2	Using Built-in DTMF Grammars.....	20
5.3	Retrieving Results.....	20
6	Monitoring Usage Details	21
6.1	Log Usage	21

6.2	Update Usage	21
6.3	Dump Channels.....	22
7	Usage Examples	23
7.1	Hotel Booking.....	23
8	Sequence Diagrams.....	32
8.1	MRCpv1	32
8.2	MRCpv2	33
9	References.....	34
9.1	AWS Lex	34
9.2	Specifications.....	34

1 Overview

This guide describes how to configure and use the Amazon Web Services (AWS) Lex plugin to the UniMRCP server. The document is intended for users having a certain knowledge of AWS Lex and UniMRCP.



1.1 Installation

For installation instructions, use one of the guides below.

- RPM Package Installation (Red Hat / Cent OS)
- Deb Package Installation (Debian / Ubuntu)

1.2 Applicable Versions

Instructions provided in this guide are applicable to the following versions.



UniMRCP 1.5.0 and above
UniMRCP Lex Plugin 1.0.0 and above

2 Supported Features

This is a brief check list of the features currently supported by the UniMRCP server running with the Lex plugin.

2.1 MRCP Methods

- ✓ DEFINE-GRAMMAR
- ✓ RECOGNIZE
- ✓ START-INPUT-TIMERS
- ✓ STOP
- ✓ SET-PARAMS
- ✓ GET-PARAMS

2.2 MRCP Events

- ✓ RECOGNITION-COMPLETE
- ✓ START-OF-INPUT

2.3 MRCP Header Fields

- ✓ Input-Type
- ✓ No-Input-Timeout
- ✓ Recognition-Timeout
- ✓ Speech-Complete-Timeout
- ✓ Waveform-URI
- ✓ Media-Type
- ✓ Completion-Cause
- ✓ Confidence-Threshold
- ✓ Start-Input-Timers
- ✓ DTMF-Interdigit-Timeout
- ✓ DTMF-Term-Timeout
- ✓ DTMF-Term-Char
- ✓ Save-Waveform
- ✓ Speech-Language
- ✓ Cancel-If-Queue
- ✓ Sensitivity-Level

2.4 Grammars

- ✓ Built-in speech, event and DTMF grammars

2.5 Results

- ✓ NLSML
- ✓ JSON

3 Configuration Format

The configuration file of the Lex plugin is located in `/opt/unimrcp/conf/umslex.xml`. The configuration file is written in XML.

3.1 Document

The root element of the XML document must be `<umslex>`.

Attributes

Name	Unit	Description
license-file	File path	Specifies the license file. File name may include patterns containing '*' sign. If multiple files match the pattern, the most recent one gets used.
credentials-file	File path	Specifies the AWS credentials file to use. File name may include patterns containing '*' sign. If multiple files match the pattern, the most recent one gets used.
init-sdk	Boolean	Specifies whether to initialize AWS SDK upon loading of the plugin. Must be set to true by default. Set it to false, if another plugin using the same AWS SDK is loaded prior to this plugin.
shutdown-sdk	Boolean	Specifies whether to shut down AWS SDK upon unloading of the plugin. Must be set to true by default. Set it to false, if another plugin using the same AWS SDK is unloaded next to this plugin.

Parent

None.

Children

Name	Unit	Description
<streaming-recognition>	String	Specifies recognition parameters of streaming recognition.

<results>	String	Specifies parameters of recognition results set in RECOGNITION-COMPLETE events.
<speech-dtmf-input-detector>	String	Specifies parameters of the speech and DTMF input detector.
<utterance-manager>	String	Specifies parameters of the utterance manager.
<rdr-manager>	String	Specifies parameters of the Recognition Details Record (RDR) manager.
<monitoring-agent>	String	Specifies parameters of the monitoring manager.
<license-server>	String	Specifies parameters used to connect to the license server. The use of the license server is optional.

Example

This is an example of a bare document.

```
<umslex license-file="umspolly_*.lic" credentials-file="*.json" init-sdk="true" shutdown-
sdk="true">
</umslex>
```

3.2 Streaming Recognition

This element specifies parameters of streaming recognition.

Attributes

Name	Unit	Description
language	String	Specifies the default language to use, if not set by the client. For a list of supported languages, visit
skip-unsupported-grammars	Boolean	Specifies whether to skip or raise an error while referencing a malformed or not supported grammar.
transcription-grammar	String	Specifies the name of the built-in speech transcription grammar. The grammar can be referenced as <i>builtin:speech/transcribe</i> or <i>builtin:grammar/transcribe</i> , where <i>transcribe</i> is the default value of this parameter.

thread-pool-size	Integer	Specifies the size of the executor thread pool provided by AWS SDK.
region	String	Specifies the AWS region of the Lex Bot. https://docs.aws.amazon.com/general/latest/gr/rande.html#lex_region
bot-name	String	Specifies the name of the Lex bot. Can be overridden by client via Vendor-Specific-Parameters.
alias	String	Specifies the alias of the Lex bot. Can be overridden by client via Vendor-Specific-Parameters.

Parent

<umslex>

Children

None.

Example

This is an example of streaming recognition element.

```
<streaming-recognition
  language="en-US"
  skip-unsupported-grammars="true"
  transcription-grammar="transcribe"
  region="us-west-2"
  bot-name="BookTrip"
  alias="Dev"
/>
```

3.3 Results

This element specifies parameters of recognition results set in RECOGNITION-COMPLETE events.

Attributes

Name	Unit	Description
format	String	Specifies the format of results to be returned to the client (use "standard" for NLSML and "json" for JSON).
indent	Integer	Specifies the indent to use while composing the results.

confidence-format	String	Specifies the format of the confidence score to be returned (use "auto" for a format based on protocol version, "mrcpv2" for a float value in the range of 0..1, "mrcpv1" for an integer value in the range of 0..100)
--------------------------	--------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Parent

<umslex>

Children

None.

Example

This is an example of results element.

```
<results
  format="standard"
  indent="0"
  confidence-format="auto"
/>
```

3.4 Utterance Manager

This element specifies parameters of the utterance manager.

Attributes

Name	Unit	Description
save-waveforms	Boolean	Specifies whether to save waveforms or not.
purge-existing	Boolean	Specifies whether to delete existing records on start-up.
max-file-age	Time interval [min]	Specifies a time interval in minutes after expiration of which a waveform is deleted. Set 0 for infinite.
max-file-count	Integer	Specifies the max number of waveforms to store. If reached, the oldest waveform is deleted. Set 0 for infinite.
waveform-base-uri	String	Specifies the base URI used to compose an

		absolute waveform URI.
waveform-folder	Dir path	Specifies a folder the waveforms should be stored in.

Parent

<umslex>

Children

None.

Example

The example below defines a typical utterance manager having the default parameters set.

```
<utterance-manager
  save-waveforms="false"
  purge-existing="false"
  max-file-age="60"
  max-file-count="100"
  waveform-base-uri="http://localhost/utterances/"
  waveform-folder=""
/>
```

3.5 RDR Manager

This element specifies parameters of the Recognition Details Record (RDR) manager.

Attributes

Name	Unit	Description
save-records	Boolean	Specifies whether to save recognition details records or not.
purge-existing	Boolean	Specifies whether to delete existing records on start-up.
max-file-age	Time interval [min]	Specifies a time interval in minutes after expiration of which a record is deleted. Set 0 for infinite.
max-file-count	Integer	Specifies the max number of records to store. If reached, the oldest record is deleted. Set 0 for infinite.

record-folder	Dir path	Specifies a folder to store recognition details records in. Defaults to <code>\${UniMRCPInstallDir}/var</code> .
----------------------	----------	------------------------------------------------------------------------------------------------------------------

Parent

<umslex>

Children

None.

Example

The example below defines a typical utterance manager having the default parameters set.

```
<rdr-manager
  save-records="false"
  purge-existing="false"
  max-file-age="60"
  max-file-count="100"
  waveform-folder=""
/>
```

3.6 Monitoring Agent

This element specifies parameters of the monitoring agent.

Attributes

Name	Unit	Description
refresh-period	Time interval [sec]	Specifies a time interval in seconds used to periodically refresh usage details. See <usage-refresh-handler>.

Parent

<umslex>

Children

<usage-change-handler>
<usage-refresh-handler>

Example

The example below defines a monitoring agent with usage change and refresh handlers.

```
<monitoring-agent refresh-period="60">

  <usage-change-handler>
    <log-usage enable="true" priority="NOTICE"/>
  </usage-change-handler>

  <usage-refresh-handler>
    <dump-channels enable="true" status-file="umslex-channels.status"/>
  </usage-refresh-handler >

</monitoring-agent>
```

3.7 Usage Change Handler

This element specifies an event handler called on every usage change.

Attributes

None.

Parent

<monitoring-agent>

Children

<log-usage>
<update-usage>
<dump-channels>

Example

This is an example of the usage change event handler.

```
<usage-change-handler>
  <log-usage enable="true" priority="NOTICE"/>
  <update-usage enable="false" status-file="umslex-usage.status"/>
  <dump-channels enable="false" status-file="umslex-channels.status"/>
</usage-change-handler>
```

3.8 Usage Refresh Handler

This element specifies an event handler called periodically to update usage details.

Attributes

None.

Parent

<monitoring-agent>

Children

<log-usage>

<update-usage>

<dump-channels>

Example

This is an example of the usage change event handler.

```
<usage-refresh-handler>  
  <log-usage enable="true" priority="NOTICE"/>  
  <update-usage enable="false" status-file="umslex-usage.status"/>  
  <dump-channels enable="false" status-file="umslex-channels.status"/>  
</usage-refresh-handler>
```

3.9 License Server

This element specifies parameters used to connect to the license server.

Attributes

Name	Unit	Description
enable	Boolean	Specifies whether the use of license server is enabled or not. If enabled, the license-file attribute is not honored.
server-address	String	Specifies the IP address or host name of the license server.
certificate-file	File path	Specifies the client certificate used to connect to the license server. File name may include patterns containing a '*' sign. If multiple files match the pattern, the most recent one gets used.
ca-file	File path	Specifies the certificate authority used to validate the license server.
channel-count	Integer	Specifies the number of channels to check out from the license server. If not specified

		or set to 0, either all available channels or a pool of channels will be checked based on the configuration of the license server.
--	--	------------------------------------------------------------------------------------------------------------------------------------

Parent

<umslex>

Children

None.

Example

The example below defines a typical configuration which can be used to connect to a license server located, for example, at 10.0.0.1.

```
<license-server
  enable="true"
  server-address="10.0.0.1"
  certificate-file="unilic_client_*.cert"
  ca-file="unilic_ca.crt"
/>
```

For further reference to the license server, visit

<http://unimrcp.org/licserver>

4 Configuration Steps

This section outlines common configuration steps.

4.1 Using Default Configuration

The default configuration should be sufficient for the general use.

4.2 Using with Polly

This section must be skipped if the Lex plugin is used without the Polly plugin. However, in case both Polly and Lex plugins are loaded into the same instance of UniMRCP server, then the plugins need to be configured in a certain way to ensure the AWS SDK is initialized and shutdown only once.

```
<umspolly license-file="umspolly_*.lic" credentials-file="aws.credentials" init-sdk="true"
shutdown-sdk="false">
<umslex license-file="umslex_*.lic" credentials-file="aws.credentials" init-sdk="false"
shutdown-sdk="true">
```

4.3 Specifying Lex Bot

The parameters of the Lex bot *region*, *bot-name* and *alias* must be specified in the configuration file *umslex.xml*. For example:

```
<streaming-recognition
  language="en-US"
  region="us-west-2"
  bot-name="BookTrip"
  alias="Dev"
/>
```

The parameters *bot-name* and *alias* can also be specified per individual MRCP SET-PARAMS or RECOGNIZE requests via the header field Vendor-Specific-Parameters.

4.4 Specifying Recognition Language

Recognition language can be specified by the client per MRCP session by means of the header field *Speech-Language* set in a SET-PARAMS or RECOGNIZE request. Otherwise, the parameter *language* set in the configuration file *umslex.xml* is used. The parameter defaults to *en-US*.

4.5 Specifying Sampling Rate

Sampling rate is determined based on the SDP negotiation. Refer to the configuration guide of the UniMRCP server on how to specify supported encodings and sampling rates to be used in communication between the client and server.

The native sampling rate with the linear16 audio encoding is used to post audio data to the Lex service.

4.6 Specifying Speech Input Parameters

While the default parameters specified for the speech input detector are sufficient for the general use, various parameters can be adjusted to better suit a particular requirement.

- `speech-start-timeout`

This parameter is used to trigger a start of speech input. The shorter is the timeout, the sooner a *START-OF-INPUT* event is delivered to the client. However, a short timeout may also lead to a false positive.

- `speech-complete-timeout`

This parameter is used to trigger an end of speech input. The shorter is the timeout, the shorter is the response time. However, a short timeout may also lead to a false positive.

- `vad-mode`

This parameter is used to specify an operating mode of the Voice Activity Detector (VAD) within an integer range of [0 ... 3]. A higher mode is more aggressive and, as a result, is more restrictive in reporting speech. The parameter can be overridden per MRCP session by setting the header field *Sensitivity-Level* in a *SET-PARAMS* or *RECOGNIZE* request. The following table shows how the *Sensitivity-Level* is mapped to the *vad-mode*.

Sensitivity-Level	Vad-Mode
[0.00 ... 0.25)	0
[0.25 ... 0.50)	1
[0.50 ... 0.75)	2
[0.75 ... 1.00]	3

4.7 Specifying DTMF Input Parameters

While the default parameters specified for the DTMF input detector are sufficient for the general use, various parameters can be adjusted to better suit a particular requirement.

- `dtmf-interdigit-timeout`

This parameter is used to set an inter-digit timeout on DTMF input. The parameter can be overridden per MRCP session by setting the header field *DTMF-Interdigit-Timeout* in a *SET-PARAMS* or *RECOGNIZE* request.

- `dtmf-term-timeout`

This parameter is used to set a termination timeout on DTMF input and is in effect when *dtmf-term-char* is set and there is a match for an input grammar. The parameter can be overridden per MRCP session by setting the header field *DTMF-Term-Timeout* in a *SET-PARAMS* or *RECOGNIZE* request.

- *dtmf-term-char*

This parameter is used to set a character terminating DTMF input. The parameter can be overridden per MRCP session by setting the header field *DTMF-Term-Char* in a *SET-PARAMS* or *RECOGNIZE* request.

4.8 Specifying No-Input and Recognition Timeouts

- *noinput-timeout*

This parameter is used to trigger a no-input event. The parameter can be overridden per MRCP session by setting the header field *No-Input-Timeout* in a *SET-PARAMS* or *RECOGNIZE* request.

- *input-timeout*

This parameter is used to limit input (recognition) time. The parameter can be overridden per MRCP session by setting the header field *Recognition-Timeout* in a *SET-PARAMS* or *RECOGNIZE* request.

4.9 Maintaining Utterances

Saving of utterances is not required for regular operation and is disabled by default. However, enabling this functionality allows to save utterances sent to the Lex service and later listen to them offline.

The relevant settings can be specified via the element *utterance-manager*.

- *save-waveforms*

Utterances can optionally be recorded and stored if the configuration parameter *save-waveforms* is set to true. The parameter can be overridden per MRCP session by setting the header field *Save-Waveforms* in a *SET-PARAMS* or *RECOGNIZE* request.

- *purge-existing*

This parameter specifies whether to delete existing waveforms on start-up.

- *max-file-age*

This parameter specifies a time interval in minutes after expiration of which a waveform is deleted. If set to 0, there is no expiration time specified.

- *max-file-count*

This parameter specifies the maximum number of waveforms to store. If the specified number is reached, the oldest waveform is deleted. If set to 0, there is no limit specified.

- *waveform-base-uri*

This parameter specifies the base URI used to compose an absolute waveform URI returned in the header field *Waveform-Uri* in response to a *RECOGNIZE* request.

- waveform-folder

This parameter specifies a path to the directory used to store waveforms in. The directory defaults to `${UniMRCPIInstallDir}/var`.

4.10 Maintaining Recognition Details Records

Producing of recognition details records (RDR) is not required for regular operation and is disabled by default. However, enabling this functionality allows to store details of each recognition attempt in a separate file and analyze them later offline. The RDRs are stored in the JSON format.

The relevant settings can be specified via the element `rdr-manager`.

- save-records

This parameter specifies whether to save recognition details records or not.

- purge-existing

This parameter specifies whether to delete existing records on start-up.

- max-file-age

This parameter specifies a time interval in minutes after expiration of which a record is deleted. If set to 0, there is no expiration time specified.

- max-file-count

This parameter specifies the maximum number of records to store. If the specified number is reached, the oldest record is deleted. If set to 0, there is no limit specified.

- record-folder

This parameter specifies a path to the directory used to store records in. The directory defaults to `${UniMRCPIInstallDir}/var`.

5 Recognition Grammars and Results

5.1 Using Built-in Speech Grammar

A pre-set built-in speech grammar can be referenced by the MRCP client in a RECOGNIZE request as follows:

```
builtin:speech/transcribe
```

5.2 Using Built-in DTMF Grammars

Pre-set built-in DTMF grammars can be referenced by the MRCP client in a RECOGNIZE request as follows:

```
builtin:dtmf/$id
```

As a result, captured DTMFs will be posted to the Lex service for intent detection.

5.3 Retrieving Results

Results received from the Lex service are transformed to one of the following formats

- NLSML
- JSON

and sent to the MRCP client in a *RECOGNITION-COMplete* event. Recognition results settings can be specified in the configuration file in the element *results*.

In case of NLSML results, the *<instance>* element contains an XML representation of the results received from the Lex service.

6 Monitoring Usage Details

The number of in-use and total licensed channels can be monitored in several alternate ways. There is a set of actions which can take place on certain events. The behavior is configurable via the element *monitoring-agent*, which contains two event handlers: *usage-change-handler* and *usage-refresh-handler*.

While the *usage-change-handler* is invoked on every acquisition and release of a licensed channel, the *usage-refresh-handler* is invoked periodically on expiration of a timeout specified by the attribute *refresh-period*.

The following actions can be specified for either of the two handlers.

6.1 Log Usage

The action *log-usage* logs the following data in the order specified.

- The number of currently in-use channels.
- The maximum number of channels used concurrently.
- The total number of licensed channels.

The following is a sample log statement, indicating 0 in-use, 0 max-used and 2 total channels.

```
[NOTICE] Lex Usage: 0/0/2
```

6.2 Update Usage

The action *update-usage* writes the following data to a status file *umslex-usage.status*, located by default in the directory *\${UniMRCPIInstallDir}/var/status*.

- The number of currently in-use channels.
- The maximum number of channels used concurrently.
- The total number of licensed channels.
- The current status of the license permit.

The following is a sample content of the status file.

```
in-use channels: 0  
max used channels: 0  
total channels: 2  
license permit: true
```

6.3 Dump Channels

The action *dump-channels* writes the identifiers of in-use channels to a status file *umslex-channels.status*, located by default in the directory `${UniMRCPIInstallDir}/var/status`.

7 Usage Examples

7.1 Hotel Booking

This example demonstrates an MRCP message exchange based on a conversation with the sample BookTrip Lex bot.

Interaction 1

Input: Book a hotel

C->S:

```
MRCP/2.0 361 RECOGNIZE 1
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Content-Id: request1@form-level
Content-Type: text/uri-list
Cancel-If-Queue: false
No-Input-Timeout: 50000
Recognition-Timeout: 10000
Start-Input-Timers: true
Confidence-Threshold: 0.87
Sensitivity-Level: 0.5
Save-Waveform: true
Content-Length: 25

builtin:speech/transcribe
```

S->C:

```
MRCP/2.0 83 1 200 IN-PROGRESS
Channel-Identifier: 2a203b9adebe40e7@speechrecog
```

S->C:

```
MRCP/2.0 115 START-OF-INPUT 1 IN-PROGRESS
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Input-Type: speech
```

S->C:

MRCP/2.0 823 RECOGNITION-COMPLETE 1 COMPLETE
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Completion-Cause: 000 success
Waveform-Uri: <http://localhost/utterances/umslex-2a203b9adebe40e7-1.wav>;size=21440;duration=1340
Content-Type: application/x-nlsml
Content-Length: 538

```
<?xml version="1.0"?>
<result>
  <interpretation grammar="builtin:speech/transcribe" confidence="1">
    <instance>
      <intent>BookHotel</intent>
      <slots>
        <CheckInDate></CheckInDate>
        <Location></Location>
        <Nights></Nights>
        <RoomType></RoomType>
      </slots>
      <message>What city will you be staying in?</message>
      <dialogstate>ElicitSlot</dialogstate>
      <slottoelicit>Location</slottoelicit>
    </instance>
    <input mode="speech">book a hotel</input>
  </interpretation>
</result>
```

Interaction 2

Input: Sunnyvale

C->S:

MRCP/2.0 361 RECOGNIZE 2
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Content-Id: request1@form-level
Content-Type: text/uri-list
Cancel-If-Queue: false
No-Input-Timeout: 50000
Recognition-Timeout: 10000
Start-Input-Timers: true
Confidence-Threshold: 0.87
Sensitivity-Level: 0.5
Save-Waveform: true
Content-Length: 25

builtin:speech/transcribe

S->C:

MRCP/2.0 83 2 200 IN-PROGRESS
Channel-Identifier: 2a203b9adebe40e7@speechrecog

S->C:

MRCP/2.0 115 START-OF-INPUT 2 IN-PROGRESS
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Input-Type: speech

S->C:

MRCP/2.0 832 RECOGNITION-COMPLETE 2 COMPLETE
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Completion-Cause: 000 success
Waveform-Uri: <http://localhost/utterances/umslex-2a203b9adebe40e7-2.wav>;size=18400;duration=1150
Content-Type: application/x-nlsml
Content-Length: 547

```
<?xml version="1.0"?>
<result>
  <interpretation grammar="builtin:speech/transcribe" confidence="1">
    <instance>
      <intent>BookHotel</intent>
      <slots>
        <CheckInDate></CheckInDate>
        <Location>sunnyvale</Location>
        <Nights></Nights>
        <RoomType></RoomType>
      </slots>
      <message>What day do you want to check in?</message>
      <dialogstate>ElicitSlot</dialogstate>
      <slottoelicit>CheckInDate</slottoelicit>
    </instance>
    <input mode="speech">sunnyvale</input>
  </interpretation>
</result>
```

Interaction 3

Input: October 1st

C->S:

```
MRCP/2.0 361 RECOGNIZE 3
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Content-Id: request1@form-level
Content-Type: text/uri-list
Cancel-If-Queue: false
No-Input-Timeout: 50000
Recognition-Timeout: 10000
Start-Input-Timers: true
Confidence-Threshold: 0.87
Sensitivity-Level: 0.5
Save-Waveform: true
Content-Length: 25

builtin:speech/transcribe
```

S->C:

```
MRCP/2.0 83 3 200 IN-PROGRESS
Channel-Identifier: 2a203b9adebe40e7@speechrecog
```

S->C:

```
MRCP/2.0 115 START-OF-INPUT 3 IN-PROGRESS
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Input-Type: speech
```

S->C:

```
MRCP/2.0 844 RECOGNITION-COMPLETE 3 COMPLETE
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Completion-Cause: 000 success
Waveform-Uri: <http://localhost/utterances/umslex-2a203b9adebe40e7-
3.wav>;size=23520;duration=1470
Content-Type: application/x-nlsml
Content-Length: 559

<?xml version="1.0"?>
<result>
```

```
<interpretation grammar="builtin:speech/transcribe" confidence="1">
  <instance>
    <intent>BookHotel</intent>
    <slots>
      <CheckInDate>2018-10-01</CheckInDate>
      <Location>sunnyvale</Location>
      <Nights></Nights>
      <RoomType></RoomType>
    </slots>
    <message>How many nights will you be staying?</message>
    <dialogstate>ElicitSlot</dialogstate>
    <slottoelicit>Nights</slottoelicit>
  </instance>
  <input mode="speech">october first</input>
</interpretation>
</result>
```

Interaction 4

Input: 2 [nights]

C->S:

```
MRCP/2.0 361 RECOGNIZE 4
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Content-Id: request1@form-level
Content-Type: text/uri-list
Cancel-If-Queue: false
No-Input-Timeout: 50000
Recognition-Timeout: 10000
Start-Input-Timers: true
Confidence-Threshold: 0.87
Sensitivity-Level: 0.5
Save-Waveform: true
Content-Length: 25

builtin:speech/transcribe
```

S->C:

```
MRCP/2.0 83 4 200 IN-PROGRESS
Channel-Identifier: 2a203b9adebe40e7@speechrecog
```

S->C:

MRCP/2.0 115 START-OF-INPUT 4 IN-PROGRESS
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Input-Type: speech

S->C:

MRCP/2.0 856 RECOGNITION-COMplete 4 COMPLETE
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Completion-Cause: 000 success
Waveform-Uri: <http://localhost/utterances/umslex-2a203b9adebe40e7-4.wav>;size=13760;duration=860
Content-Type: application/x-nlsml
Content-Length: 572

```
<?xml version="1.0"?>
<result>
  <interpretation grammar="builtin:speech/transcribe" confidence="1">
    <instance>
      <intent>BookHotel</intent>
      <slots>
        <CheckInDate>2018-10-01</CheckInDate>
        <Location>sunnyvale</Location>
        <Nights>2</Nights>
        <RoomType></RoomType>
      </slots>
      <message>What type of room would you like, queen, king or deluxe?</message>
      <dialogstate>ElicitSlot</dialogstate>
      <slottoelicit>RoomType</slottoelicit>
    </instance>
    <input mode="speech">two</input>
  </interpretation>
</result>
```

Interaction 5

Input: King

C->S:

MRCP/2.0 361 RECOGNIZE 5
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Content-Id: request1@form-level
Content-Type: text/uri-list

Cancel-If-Queue: false
No-Input-Timeout: 50000
Recognition-Timeout: 10000
Start-Input-Timers: true
Confidence-Threshold: 0.87
Sensitivity-Level: 0.5
Save-Waveform: true
Content-Length: 25

builtin:speech/transcribe

S->C:

MRCP/2.0 83 5 200 IN-PROGRESS
Channel-Identifier: 2a203b9adebe40e7@speechrecog

S->C:

MRCP/2.0 115 START-OF-INPUT 5 IN-PROGRESS
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Input-Type: speech

S->C:

MRCP/2.0 869 RECOGNITION-COMplete 5 COMPLETE
Channel-Identifier: 2a203b9adebe40e7@speechrecog
Completion-Cause: 000 success
Waveform-Uri: <http://localhost/utterances/umslex-2a203b9adebe40e7-5.wav>;size=12960;duration=810
Content-Type: application/x-nlsml
Content-Length: 585

```
<?xml version="1.0"?>
<result>
  <interpretation grammar="builtin:speech/transcribe" confidence="1">
    <instance>
      <intent>BookHotel</intent>
      <slots>
        <CheckInDate>2018-10-01</CheckInDate>
        <Location>sunnyvale</Location>
        <Nights>2</Nights>
        <RoomType>king</RoomType>
      </slots>
    </instance>
  </interpretation>
</result>
```

```
<message>Okay, I have you down for a 2 night stay in sunnyvale starting 2018-10-01.  
Shall I book the reservation?</message>  
<dialogstate>ConfirmIntent</dialogstate>  
</instance>  
<input mode="speech">king</input>  
</interpretation>  
</result>
```

Interaction 6

Input: Yes

C->S:

```
MRCP/2.0 361 RECOGNIZE 6  
Channel-Identifier: 2a203b9adebe40e7@speechrecog  
Content-Id: request1@form-level  
Content-Type: text/uri-list  
Cancel-If-Queue: false  
No-Input-Timeout: 50000  
Recognition-Timeout: 10000  
Start-Input-Timers: true  
Confidence-Threshold: 0.87  
Sensitivity-Level: 0.5  
Save-Waveform: true  
Content-Length: 25  
  
builtin:speech/transcribe
```

S->C:

```
MRCP/2.0 83 6 200 IN-PROGRESS  
Channel-Identifier: 2a203b9adebe40e7@speechrecog
```

S->C:

```
MRCP/2.0 115 START-OF-INPUT 6 IN-PROGRESS  
Channel-Identifier: 2a203b9adebe40e7@speechrecog  
Input-Type: speech
```

S->C:

MRCP/2.0 743 RECOGNITION-COMPLETE 6 COMPLETE

Channel-Identifier: 2a203b9adebe40e7@speechrecog

Completion-Cause: 000 success

Waveform-Uri: <http://localhost/utterances/umslex-2a203b9adebe40e7-6.wav>;size=14880;duration=930

Content-Type: application/x-nlsml

Content-Length: 459

```
<?xml version="1.0"?>
```

```
<result>
```

```
  <interpretation grammar="builtin:speech/transcribe" confidence="1">
```

```
    <instance>
```

```
      <intent>BookHotel</intent>
```

```
      <slots>
```

```
        <CheckInDate>2018-10-01</CheckInDate>
```

```
        <Location>sunnyvale</Location>
```

```
        <Nights>2</Nights>
```

```
        <RoomType>king</RoomType>
```

```
      </slots>
```

```
      <dialogstate>ReadyForFulfillment</dialogstate>
```

```
    </instance>
```

```
    <input mode="speech">yes</input>
```

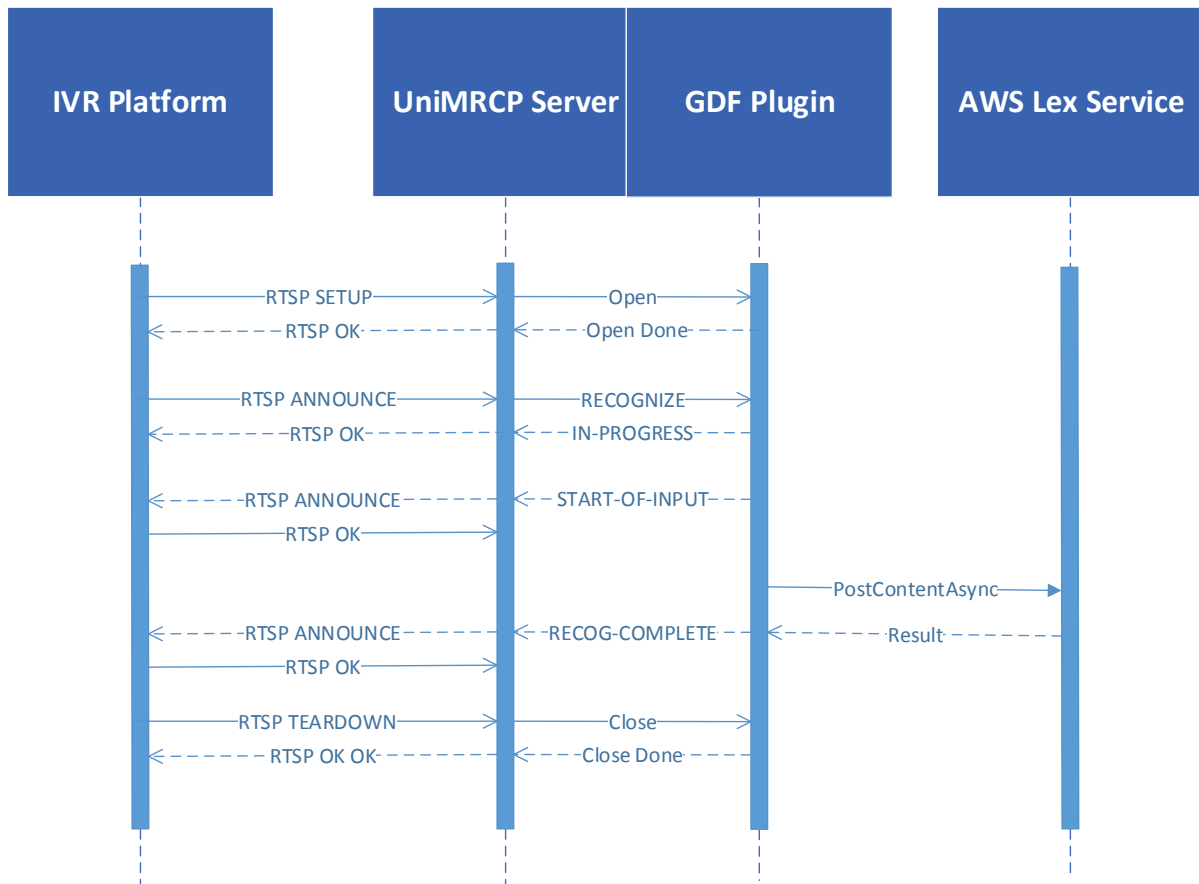
```
  </interpretation>
```

```
</result>
```

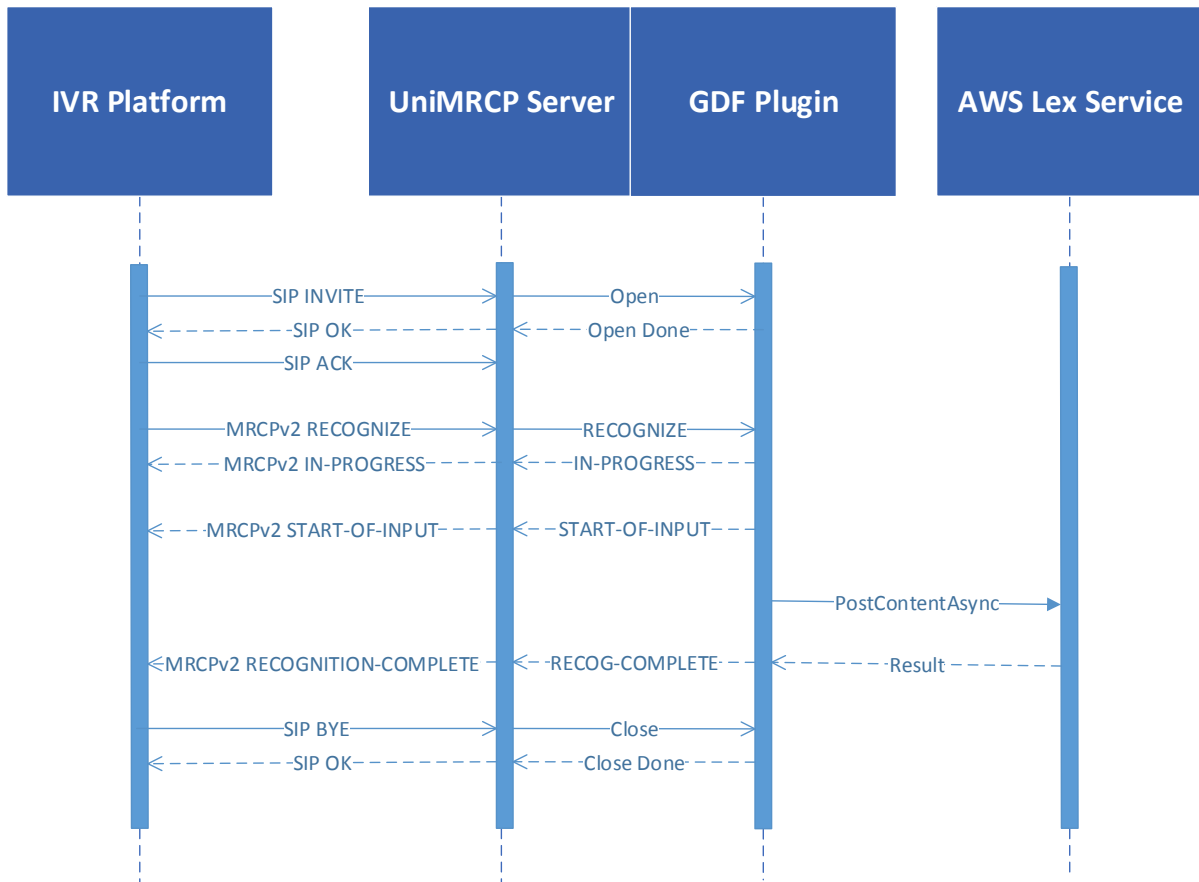
8 Sequence Diagrams

The following sequence diagrams outline common interactions between all the main components involved in a typical recognition session performed over MRCPv1 and MRCPv2 respectively.

8.1 MRCPv1



8.2 MRCPv2



9 References

9.1 AWS Lex

- [What is Amazon Lex](#)
- [How It Works](#)
- [API Reference](#)

9.2 Specifications

- [Speech Recognizer Resource](#)
- [NLSML Results](#)