Auto speech adaptation

When performing a detect intent request, you can optionally supply speech context (/dialogflow/docs/reference/rest/v2/QueryInput#SpeechContext) to provide hints to the speech recognizer. These hints can help with recognition in a specific conversation state.

The auto speech adaptation feature improves the speech recognition accuracy of your agent by automatically using agent entities, training phrases, and conversation state as speech context hints for all detect intent requests. This feature is enabled by default.

Providing explicit `speech_contexts` for a detect intent request overrides the implicit speech context hints ...ated by auto speech adaptation for input audio (speech-to-text) configuration.

Enable or disable auto speech adaptation

To enable or disable auto speech adaptation:

1. Go to the Dialogflow Console (https://dialogflow.cloud.google.com)

2. Select your agent near the top of the left sidebar menu

3. Click the settings ⚙ button next to the agent name

4. Select the **Speech** tab

5. Scroll to the **Improve Speech Recognition Quality** section

6. Toggle **Enable Auto Speech Adaptation** on or off

Example speech recognition improvements

With auto speech adaptation enabled, you can build your agent in ways to take advantage of it. The following examples show how speech recognition may be improved with certain changes to your agent:

1. If you define a training phrase that uses the `@sys.number` system entity, and the end user says "two", it may be recognized as "to", "too", "2", or "two". With auto speech adaptation

enabled, Dialogflow uses the `@sys.number` entity as a hint during speech recognition, and the parameter is recognized as "two" or "2".

2. If you define training phrases with a phrase like "stuffy nose", a similar sounding user utterance is reliably recognized as "stuffy nose" and not "stuff he knows".

3. If you define entities for product or service names offered by your company, and the end user mentions these terms in an utterance, they are more likely to be recognized.

4. User utterances with consecutive but distinct number entities can be ambiguous. For example, "I want two sixteen packs" might mean 2 quantities of 16 packs, or 216 quantities of packs. Speech Adaptation can help disambiguate these cases if you set up entities with spelled-out values:

   - Define a `quantity` entity with entries:
     ```
     zero
     one
     ...
     twenty
     ```

   - Define a `product` or `size` entity with entries:
     ```
     sixteen pack
     two ounce
     ...
     five liter
     ```

   - Only entity synonyms are used in speech adaptation, so you can define an entity with reference value `1` and single synonym `one` to simplify your fulfillment logic.

5. You can use auto speech adaptation with <u>Regexp entities</u> (/dialogflow/docs/entities-regexp). You can use these features together to capture alphanumeric or digit sequences.

   - Special characters and punctuation like `` ` ~ ! @ # $ % ^ & * ( ) - _ = + , . < > / ? ; ' : " [ ] { } \ | `` are not supported.

   - `\s*` and `\s?` are allowed, but `\s` is not.

   - You can add additional regexp entity entries that do not conform to these specificactions. You will benefit from auto speech adaptation so long as *at least one* of your entries conforms.

   For example, you can use regular expressions like the following to capture alphanumeric sequences:

```
^[A-Za-z0-9]{1,10}$
^WAC\d+$
215[2-8]{3}[A-Z]+$
[a-zA-Z]\s?[a-zA-Z]\s?[0-9]\s?[0-9]\s?[0-9]\s?[a-zA-Z]\s?[a-zA-Z]
```

The following regular expressions will adapt for digit sequences:

```
\d{2,8}
^[0-9]+$
2[0-9]{7}
[2-9]\d{2}[0-8]{3}\d{4}
```

Also consider using `@sys.number-sequence` for accepting any digit sequence, and `@sys.phone-number` for a localized phone number recognizer.

Limitations

The following limitations apply:

1. Auto speech adaptation does not work for Actions on Google (Google Assistant), because speech recognition is performed by Actions on Google before sending data to Dialogflow.

2. Auto speech adaptation's built-in support for system and regexp entities varies by language. Check Speech class tokens' (https://cloud.google.com/speech-to-text/docs/class-tokens) `$OOV_CLASS_ALPHANUMERIC_SEQUENCE` and `$OOV_CLASS_DIGIT_SEQUENCE` for supported languages. If your language is not listed, you can work around this limitation. For example, if you want an employee ID that is three letters followed by three digits to be accurately recognized, you could build your agent with the following entities and parameters:

   - Define a `digit` entity that contains 10 entity entries (with synonyms):
     ```
     0, 0
     1, 1
     ...
     9, 9
     ```

- Define a `letter` entity that contains 26 entity entries (with synonyms):
  ```
  A, A
  B, B
  ...
  Z, Z
  ```

- Define a `employee-id` entity that contains a single entity entry (without synonyms):
  `@letter @letter @letter @digit @digit @digit`

- Use `@employee-id` as a parameter in a training phrase.

★ **Note:** The `employee-id` entity entry requires whitespace between consecutive entities in order to be a valid entity definition. Because of this, a user utterance of "123ABC" is not matched, but a user utterance of "1 2 3 A B C" is matched. This is not usually a problem, because a spoken user utterance is processed with whitespace padding for alphanumeric characters.

Previous

← **Agent validation** (/dialogflow/docs/agents-validation)

Next

Data logging and enhanced speech models (/dialogflow/docs/data-logging) →