



# Microsoft Cognitive Services

Language & Speech



”

# LANGUAGE

Process text and learn how to recognize  
what users want

Bing Spell Check | Language Understanding |  
Linguistic Analysis | Text Analytics | Web Language Model |  
Translator Text and Speech

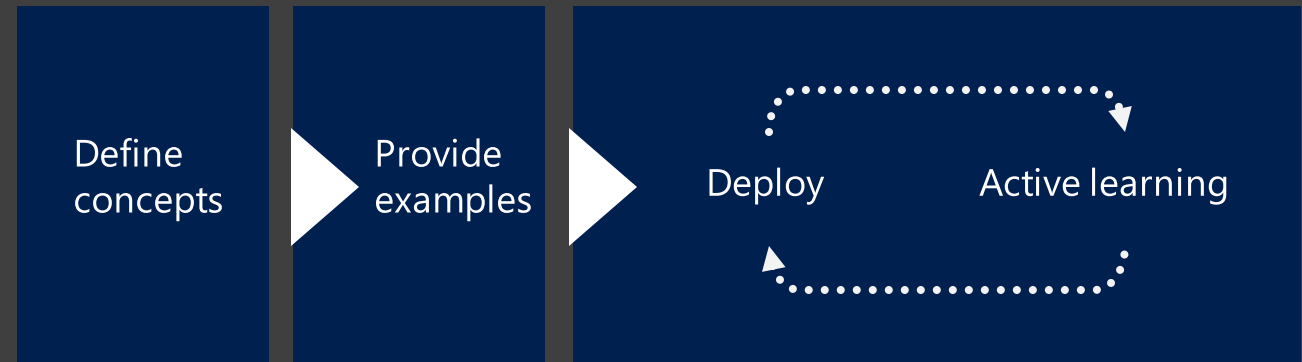
# Language Understanding Intelligent Service

Reduce labeling effort with interactive featurizing

Use visualizations to gauge performance and improvements

Leverage speech recognition with seamless integration

Deploy using just a few examples with active learning



# Language understanding models

"News about flight delays"



```
{
  "entities": [
    {
      "entity": "flight_delays",
      "type": "Topic"
    }
  ],
  "intents": [
    {
      "intent": "FindNews",
      "score": 0.99853384
    },
    {
      "intent": "None",
      "score": 0.07289317
    },
    {
      "intent": "ReadNews",
      "score": 0.0167122427
    },
    {
      "intent": "ShareNews",
      "score": 1.0919299E-06
    }
  ]
}
```





# Pizza Bot

Version: 0.1

Settings

Dashboard

Intents

Entities

Prebuilt domains PREVIEW

Features

Train & Test

Publish App

← Back to App list

## Test your application

Use this tool to test the current and published versions of your application, to check if you are progressing on the right track ... [Learn more](#)

Train Application

Last train: Jan 18, 2017 2:41:01 PM | Last publish: Jan 18, 2017 2:41:25 PM

### Interactive Testing | Batch Testing

☐ Enable published model

Labels view (Ctrl+E)

Entities



[Reset console](#)

Type a test utterance & press Enter



i ' d like a [ \$Size ] pie with [ \$Toppings ]

i need a [ \$Size ] pizza with [ \$Toppings ] and tomatoes

i need a [ \$Size ] pizza

#### Current version results

Top scoring intent

Order-Pizza (0.88)

Other intents

None (0.06) Cancel (0)

# Linguistic analysis

## Analysis tools for natural language processing

Access to part-of-speech tagging and parsing, identifying concepts, and actions





# Linguistic analysis



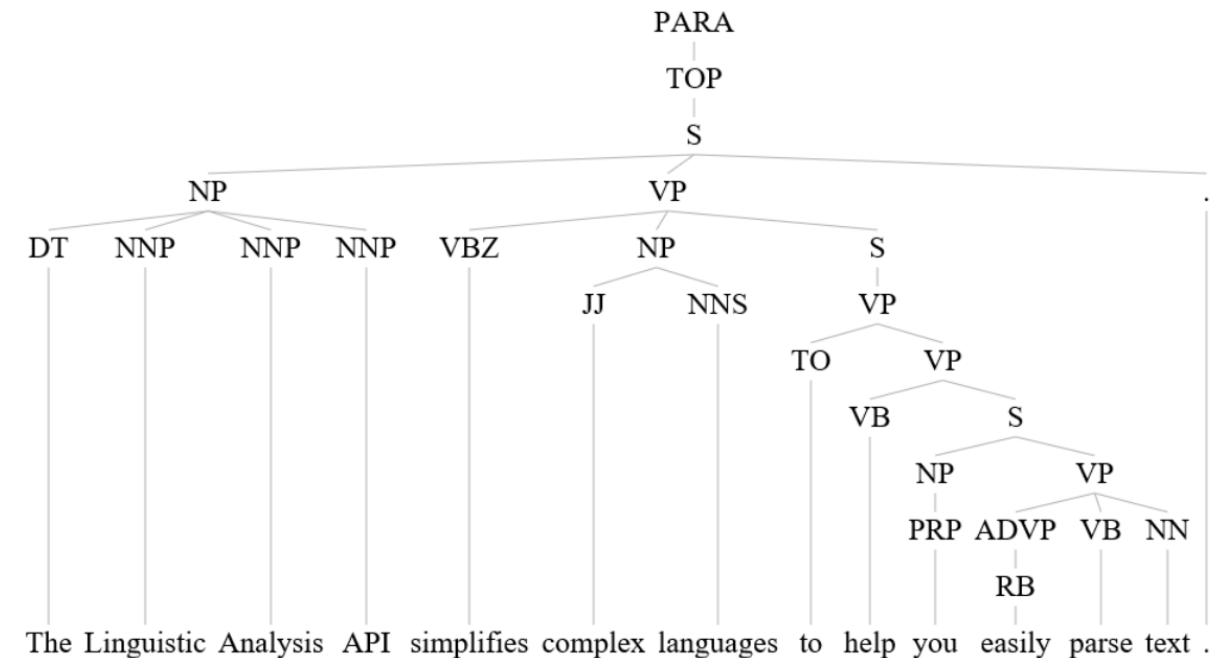
## Enter a sentence

The Linguistic Analysis API simplifies complex languages to help you easily parse text.

## POS tags

[["DT", "NNP", "NNP", "NNP", "VBZ", "JJ", "NNS", "TO", "VB", "PRP", "RB", "VBP", "NN", "."]]

## Constituency tree





# Text analytics

## Sentiment analysis

Understand if a record has positive or negative sentiment

## Key phrase extraction

Extract key phrases from a piece of text, and retrieve topics

## Language detection

Identify the language,  
120 supported languages



# Microsoft Translator

## Translator Text API

Automatically detect language  
and easily power translation to and  
from 60 supported text languages

## Translator Speech API

Easily translate real-time speech  
conversations in 9 support languages



Hear and speak to your users by filtering noise,  
identifying speakers, and understanding intent



# Speech

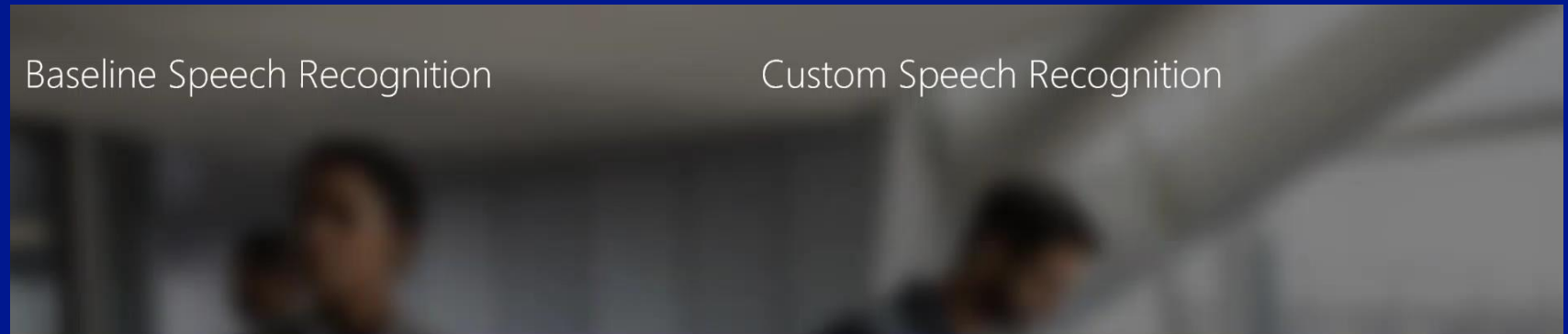
Unified Speech service |

Speech to Text | Text to Speech | Speech Translation

Speaker Recognition

Baseline Speech Recognition

Custom Speech Recognition





# Speech



## Speaker Recognition

Use speech to identify and verify individual speakers



## Translator Speech

Easily conduct real-time speech translation with a simple REST API call



## Custom Speech

Overcome speech recognition barriers like speaking style, background noise, and vocabulary



## Unified Speech service

Unified speech service for **speech-to-text** (general and custom speech models), **text-to-speech** (general and custom voice models) and **speech translation** (general and custom translator)



## Bing Speech

Convert speech to text and back again to understand user intent



# Unified Speech

Combine speech recognition, text-to-speech, customized voice models, and translation service

## **Speech to text**

Convert continuous human speech to text that can be used as inputs

## **Text to speech**

Convert text to audio files of natural-sounding speech

## **Speech translation**

Use text or speech to provide translations of speech to other languages



# Speaker Recognition

## **Speaker verification**

Check if two voices are the same

## **Speaker identification**

Identify who is speaking



# Speaker Recognition

## Enrollment

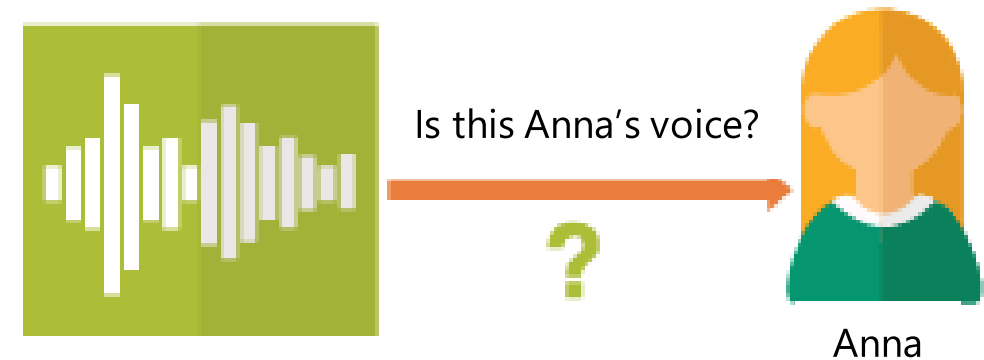
Create a unique voiceprint for a profile

## Recognition

After enrolling one or more voices, identify who is speaking from an audio clip

## Verification

Confirm if a voice belongs to a previously enrolled profile





# Bing Speech

## **Voice recognition (speech to text)**

Converts spoken audio to text

## **Voice output (text to speech)**

Synthesize audio from text

## **Speech intent recognition**

Convert spoken audio to intent



# Custom Speech

## **Customize both language and acoustic models**

Tailor speech recognition to  
your app and environment



# Custom Speech

Create custom language models for the vocabulary of the application

Adapt acoustic models to better match the expected environment of the application's users

Deploy to a custom endpoint and access from any device



Record audio



Transcribe



Adapt

Deploy





# STARSHIP COMMANDER

"The only reason we can build a product like this is because we are building on the deep learning and speech recognition expertise at Microsoft to deliver an entertainment experience that will be revolutionary."

Alexander Mejia, Owner and Creative Director,  
Human Interact

[Custom Speech Service, LUIS](#)

[Read case study here](#)

[See video here](#)



# Translator Speech

Facilitate end-to-end, real-time speech translations in over 60 languages

## **Transcribe and translate**

Optimize translations for real-life conversations

## **Customize**

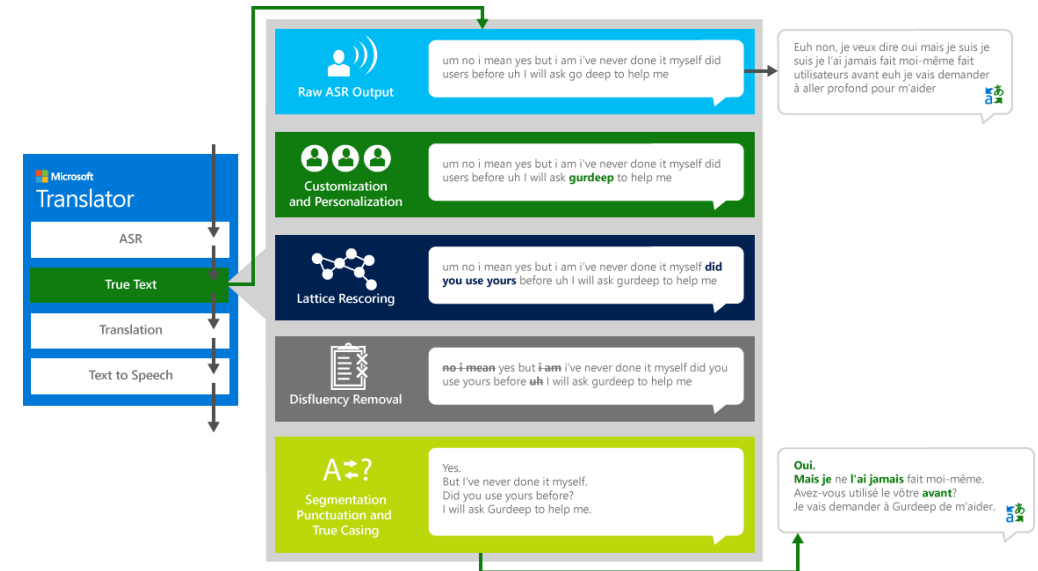
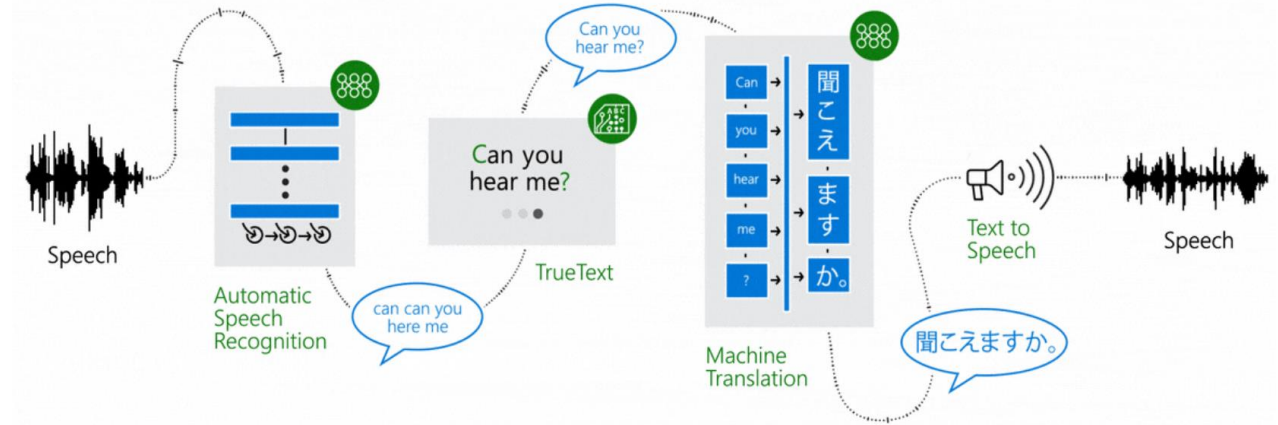
Personalize speech recognition, translations, and speech-to-text to your specific domain or scenario



# Translator Speech

To translate the "source" speech from one language to a different "target" language, Translator Speech goes through a four-step process

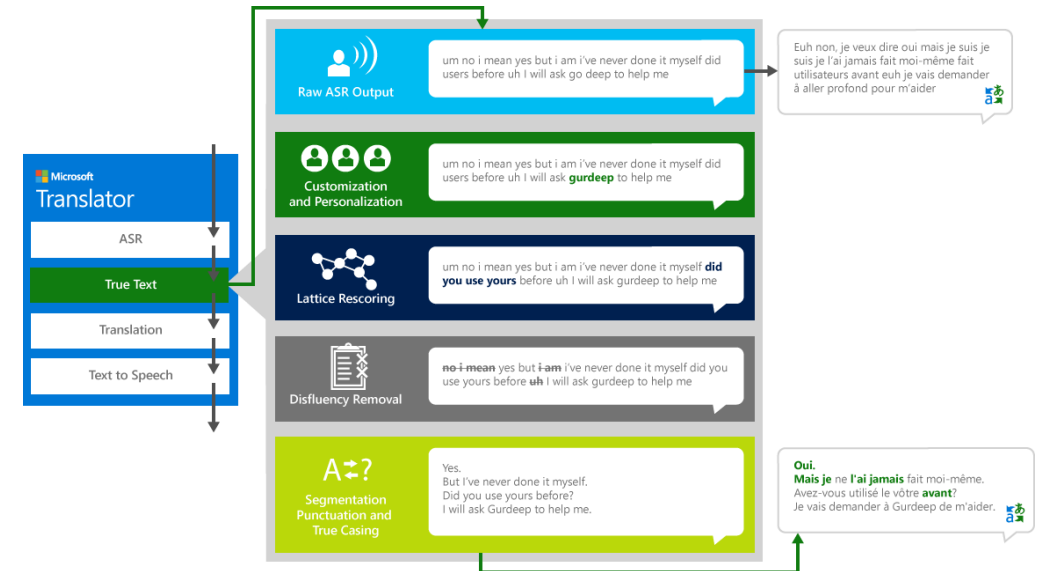
- 1 Speech recognition, to convert audio in into text
- 2 TrueText normalizes text to make it more appropriate for translation
- 3 Translation through the text translation engine with models developed for real-life spoken conversations
- 4 Text-to-speech, when necessary, to produce the translated audio



# Translator Speech

To translate the "source" speech from one language to a different "target" language, Translator Speech goes through a four-step process

- 1 Speech recognition, to convert audio in into text
- 2 TrueText normalizes text to make it more appropriate for translation
- 3 Translation through the text translation engine with models developed for real-life spoken conversations
- 4 Text-to-speech, when necessary, to produce the translated audio







# Speech & Text – Hands on Lab