

The background features a dark blue field on the left with white text. On the right, there are geometric shapes: a grey trapezoid at the top with white wavy lines, a blue trapezoid below it, and a 3D wireframe plot with a color gradient from yellow to blue. Faint blue circuit-like patterns are visible in the bottom right.

MATLAB EXPO 2017

Reusing and Prototyping Code to
Accelerate Innovation

The Case of Smart Voice Interfaces

Gabriele Bunkheila

What Device Is This?



- Innovate

```

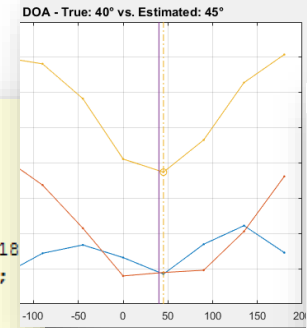
% Direction to estimate
incidentAngle = [40;0];

% Theta increments
thetaStep = 45;
steeringThetas = -180:thetaStep:180;
steeringAngles = [steeringThetas;

frameSize = 512;

s = beamformArraysSpecificationsDOA(

```



Dataset Recorder

Rec Play

Recording Subfolder: H1 Gab BlueCoin Close

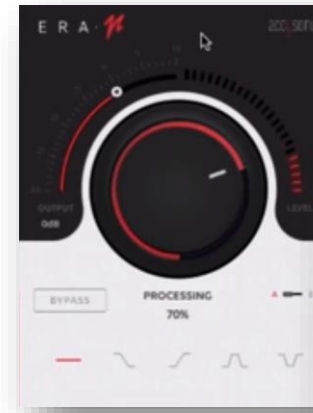
SpeechContent: the birch canoe slid on the

SpeakerName: Gabriele Bunkheila

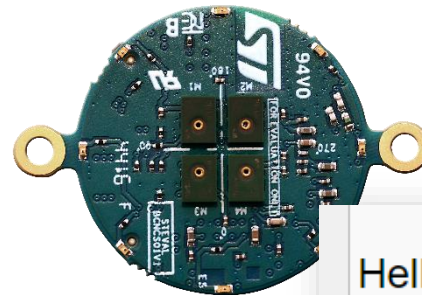
SpeakerGender: Male

SpeakerAge: 41

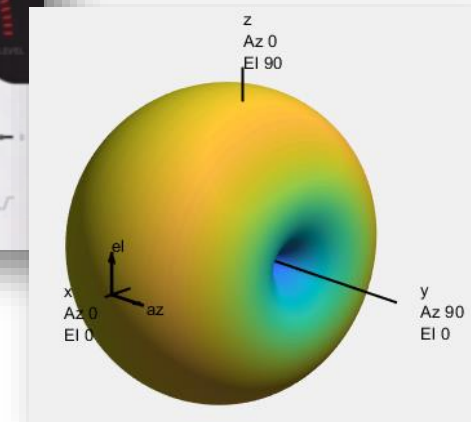
- Reuse



- Prototype



Hello
Anyone out there

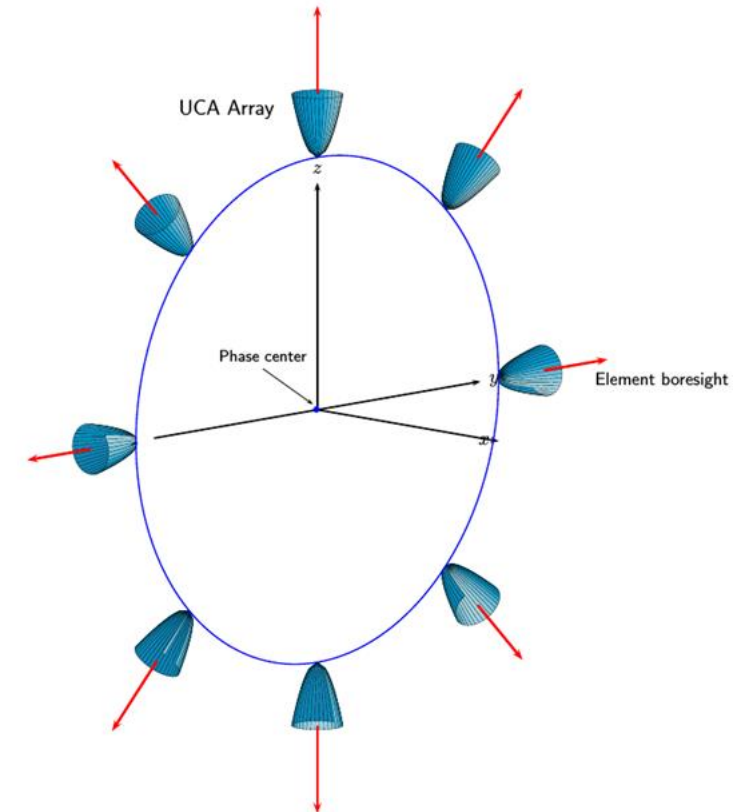
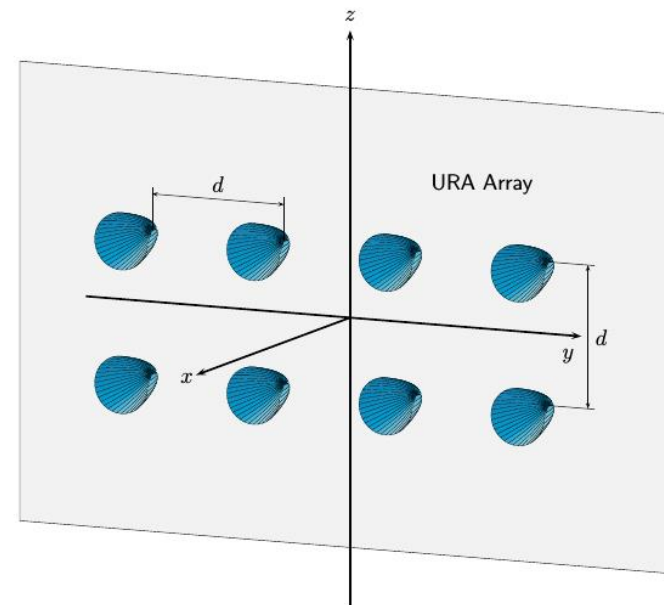
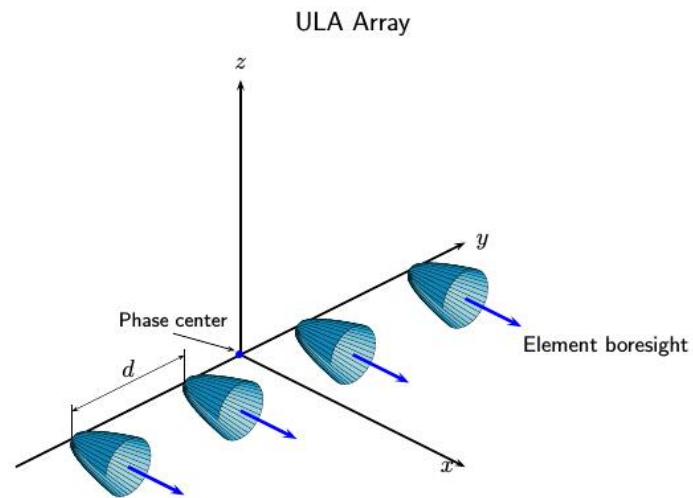


System Prototype Demo

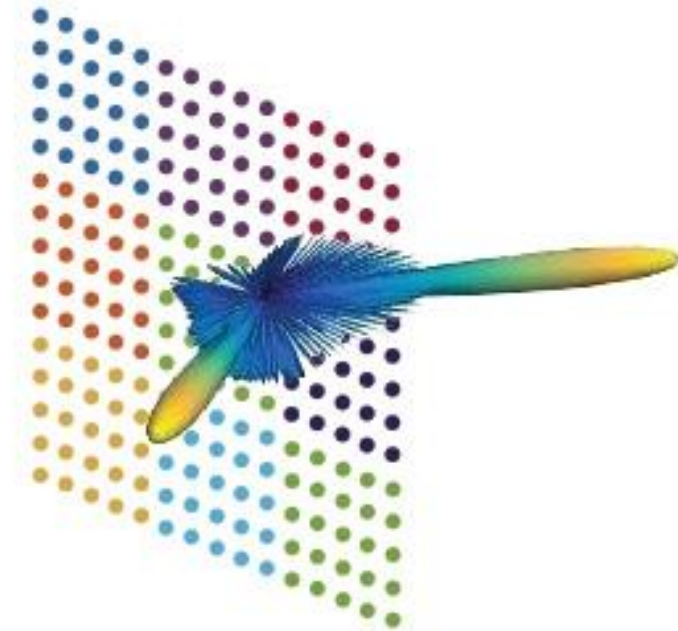
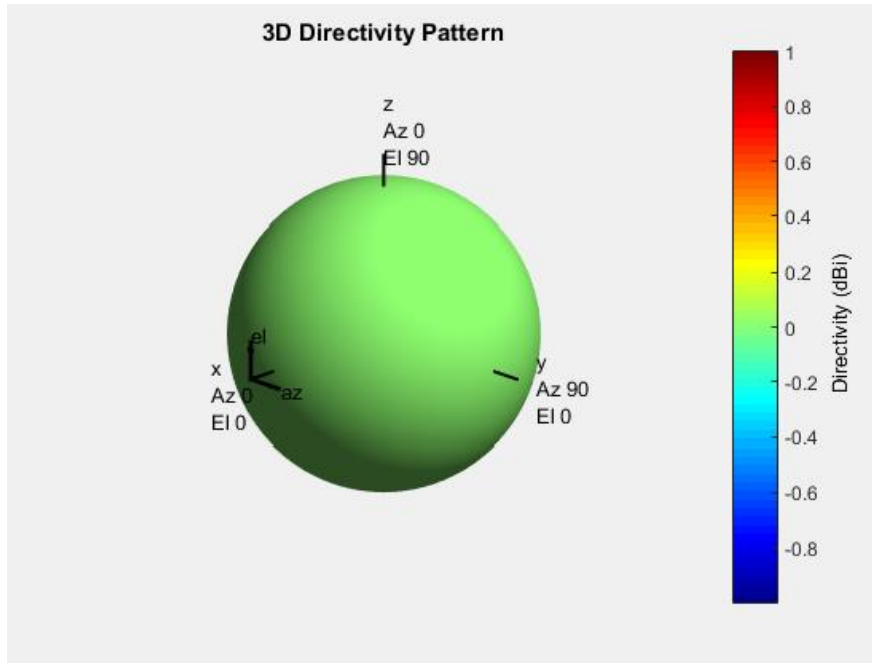
Live voice detection, beamforming and speech recognition

[Video placeholder]

What Are Microphone Arrays?

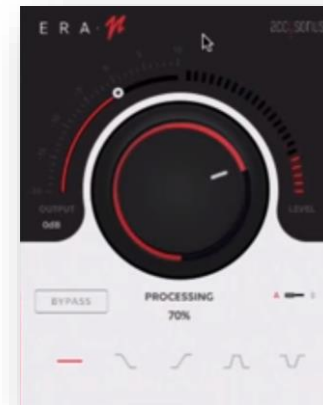
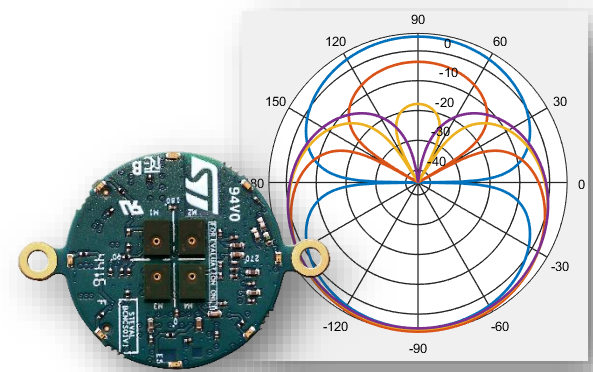


Why Microphone Arrays?



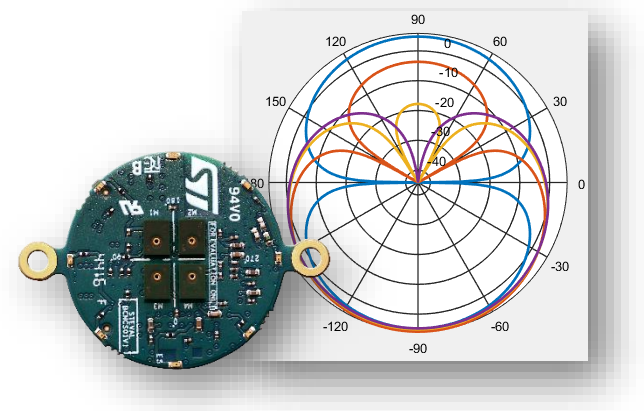
How Can I...

1. Design and simulate a microphone array system?
2. Validate my voice interface can work in real-life scenarios?
3. Understand what else can help me improve my performance?



How Can I...

1. Design and simulate a microphone array system?

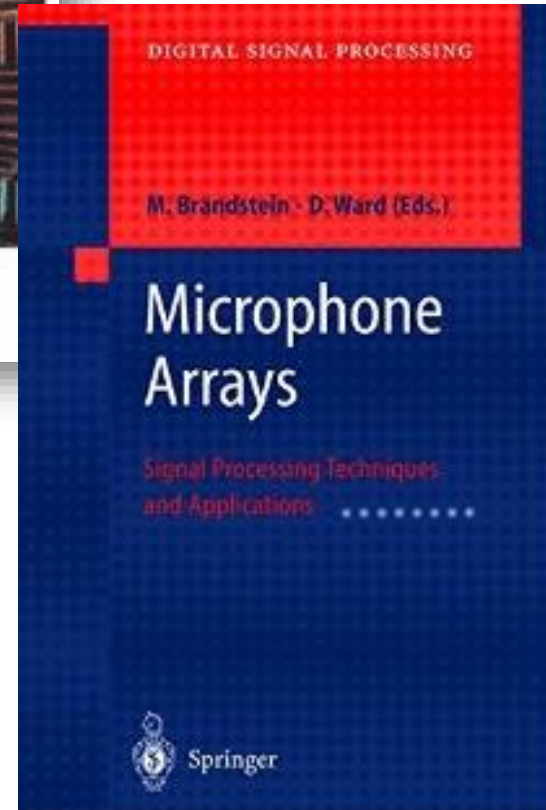
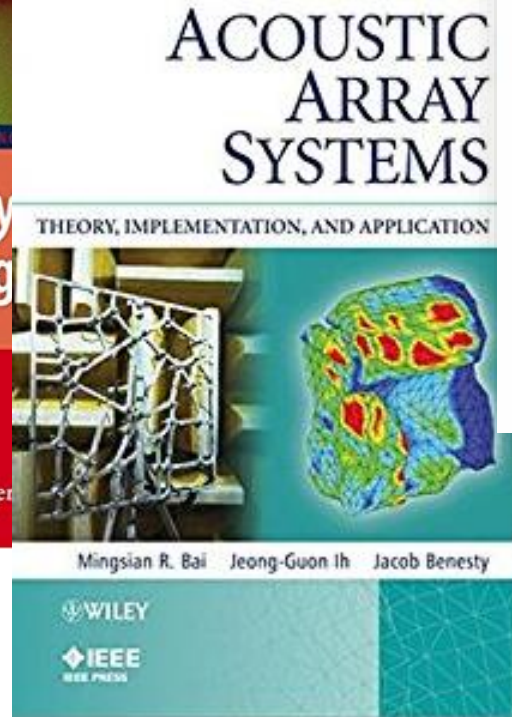
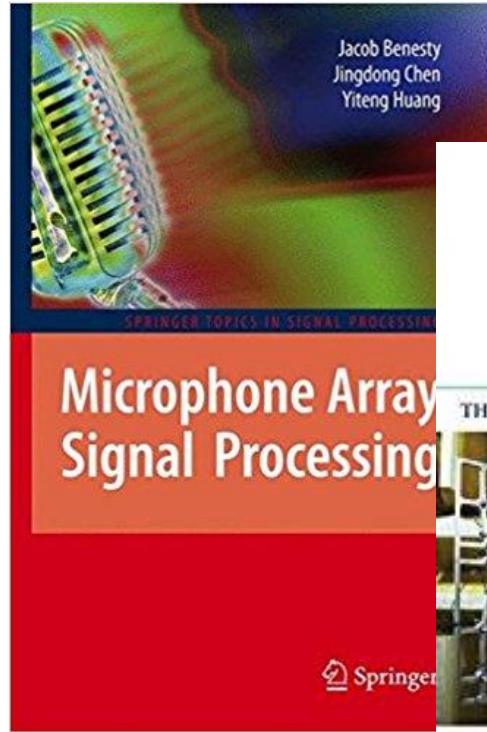


Section Overview

From analysing a microphone topology to simulating a beamforming scenario

[Video placeholder]

How To Get Started?



Demo

App-aided exploration of array topologies

[Video placeholder]

Demo

Selecting a set of directionally-selective array configurations

[Video placeholder]

Demo

Simulating time-domain sound acquisition through a microphone array

[Video placeholder]

Demo

Interactive beamforming simulation with voice source

[Video placeholder]

How Can I...

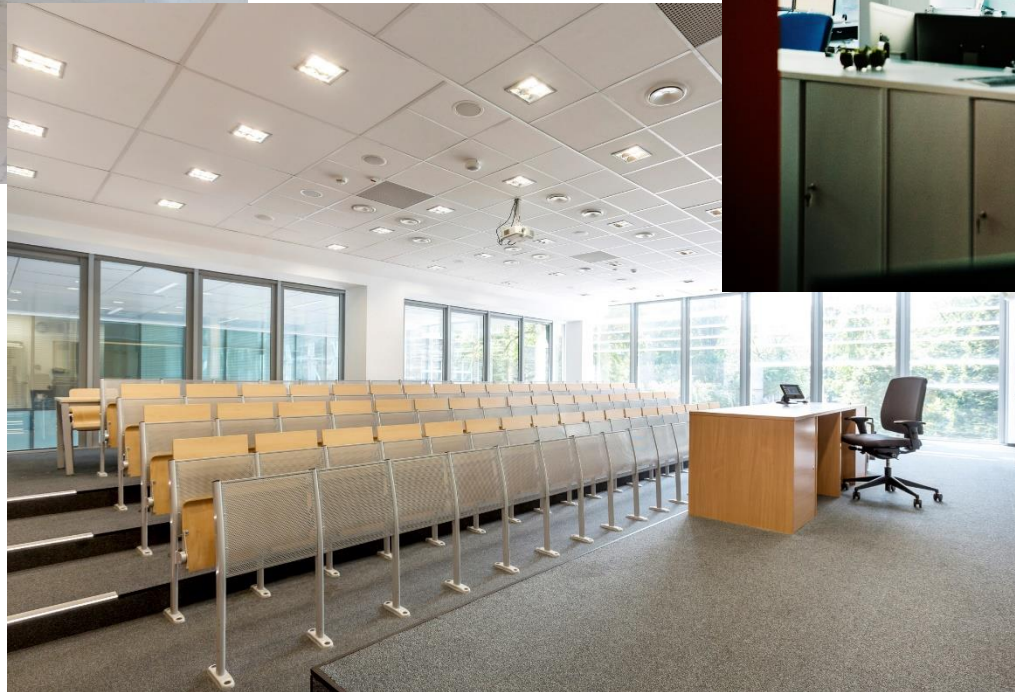
2. Validate my voice interface can work in real-life scenarios?



Constrained Simulations vs. Real Life



Constrained Simulations vs. Real Life



Demo

Streaming live multichannel audio in and out of MATLAB

[Video placeholder]

Demo

Real-time prototyping with live tuning of beamforming algorithms in MATLAB

[Video placeholder]

Demo

Common uses of Digital Audio Workstations (DAW) and audio plugins for sound acquisition and audio enhancement

[Video placeholder]

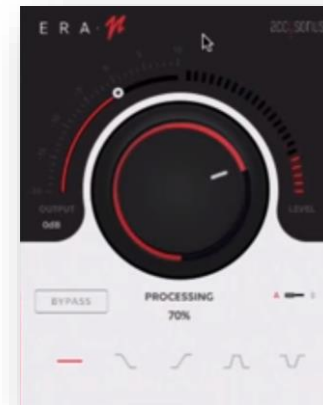
Demo

Generating VST plugins from MATLAB and prototyping array processing algorithms on Digital Audio Workstations

[Video placeholder]

How Can I...

3. Understand what else can help me improve my performance?

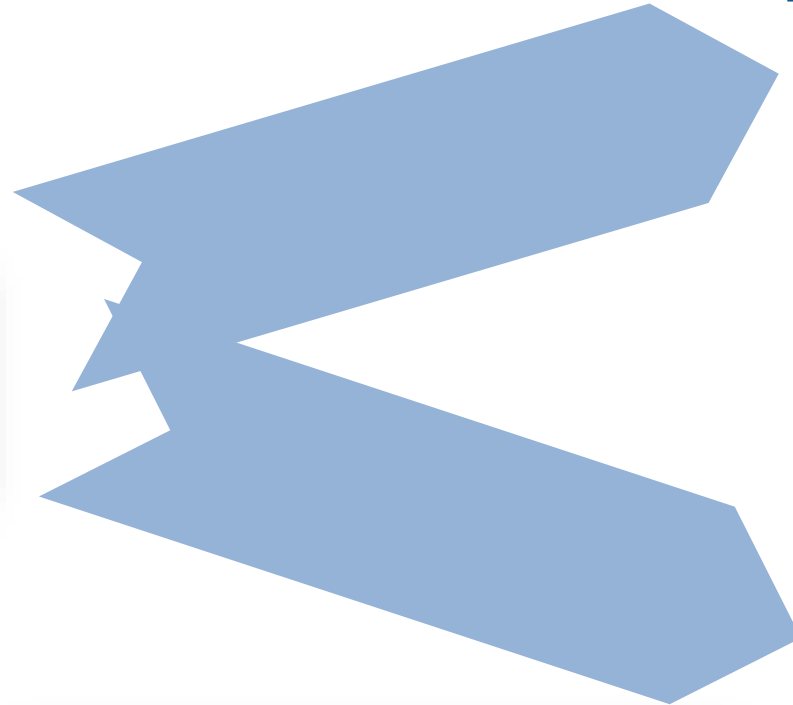
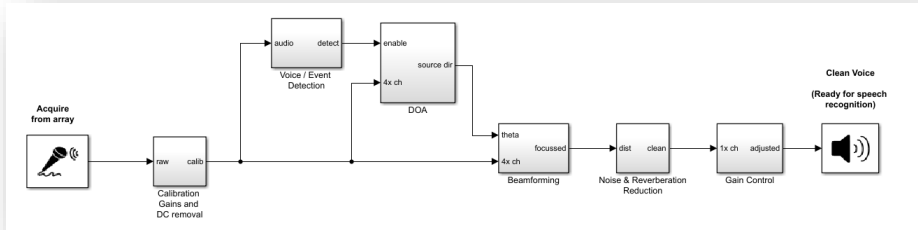


Recap Demo

Using prototypes to brainstorm on incremental improvements early on

[Video placeholder]

How To Measure Performance?

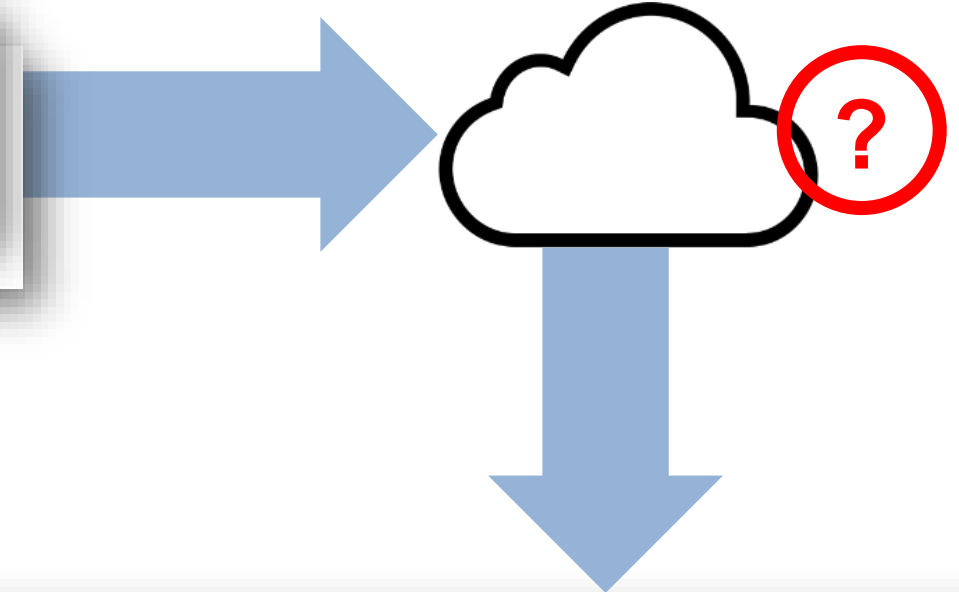
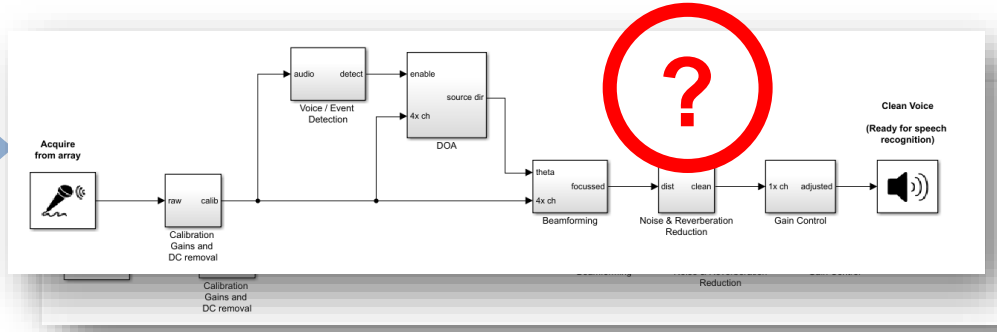


- Output audio "sounds good"



- "91.5% of spoken sentences correctly converted"

True	Guessed
e slid on the smooth planks"	"the birth canal we sl
in stockings is hard to sell"	"a large size in stock
to the dark blue background"	"blue the sheet to the
ll the depth of a well"	"it's easy to tell the
hicken leg is a rare dish"	"these days a chicken
served in round bowls"	"rice is often served



Current Folder

Name	Size
Folder	
Harvard H1	
H1 Gab BlueCoin...	
1.wav	231 KB
2.wav	217 KB
3.wav	209 KB
4.wav	249 KB
5.wav	269 KB
6.wav	251 KB
7.wav	253 KB
8.wav	269 KB
9.wav	277 KB
10.wav	269 KB
H1 Gab BlueCoin...	
1.wav	241 KB
2.wav	227 KB
3.wav	239 KB
4.wav	293 KB
5.wav	269 KB
6.wav	253 KB
7.wav	321 KB
8.wav	245 KB
9.wav	311 KB
10.wav	287 KB
H1 Gab Headset	
1.wav	253 KB
2.wav	263 KB

True	GussedBaseline	ConfidenceBaselin
"the birch canoe slid on the smooth planks"	"the birth canal we slid on the smooth planks"	0.74933
" a large size in stockings is hard to sell"	"a large size in stockings is hard to sell"	0.9366
"glue the sheet to the dark blue background"	"blue the sheet to the dark blue background"	0.95009
"its easy to tell the depth of a well"	"it's easy to tell the depth of a well"	0.97403
"these days a chicken leg is a rare dish"	"these days a chicken leg is a redfish"	0.86686
"rice is often served in round bowls"	"rice is often served in rum balls"	0.95946
"the juice of lemons makes fine punch"	"the cheese of lemons makes flying punch"	0.79089
"the box was thrown beside the parked truck"	"the box was thrown beside the box truck"	0.82796
"the hogs were fed chopped corn and garbage"	"the Hawks were fed chops corn and garbage"	0.81863
"four hours of steady work faced us"	"4 hours of study world fastest"	0.82847

Getting Started With a Cloud Service for Speech Recognition

[Video placeholder]

Hide Complexity to Accelerate Reuse

```
>> [samples, fs] = audioread('hellomatlabexpo.wav');  
>> soundsc(samples, fs)  
>> [speechString, confidenceInterval] = speech2text(samples, fs , 'en-US')
```

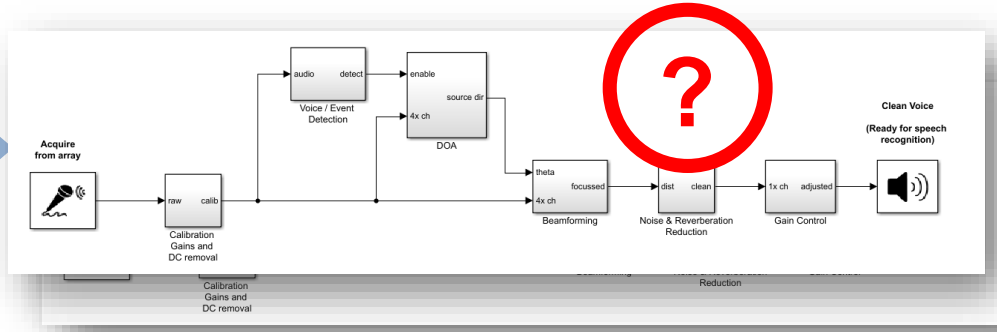


```
speechString =
```

```
    'hello Matlab Expo'
```

```
confidenceInterval =
```

```
    0.9526
```

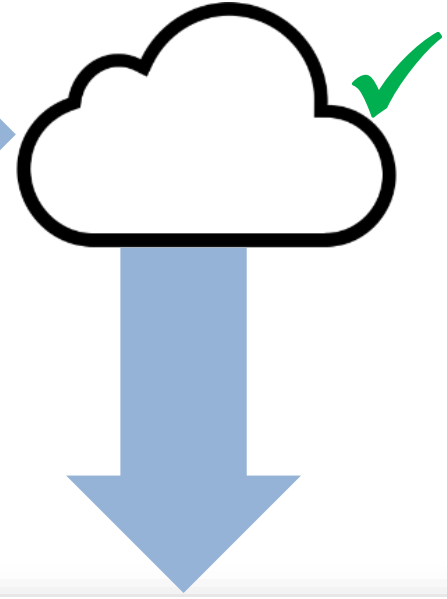


Current Folder

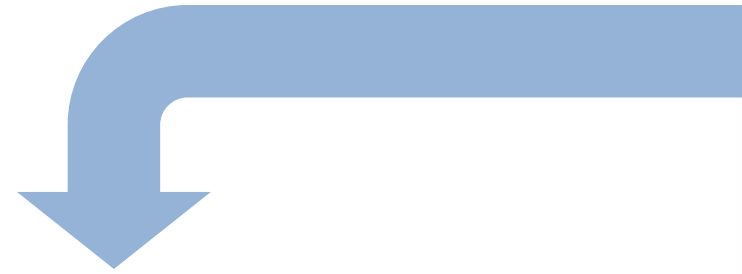
Name	Size
Folder	
Harvard H1	
H1 Gab BlueCoin...	
1.wav	231 KB
2.wav	217 KB
3.wav	209 KB
4.wav	249 KB
5.wav	269 KB
6.wav	251 KB
7.wav	253 KB
8.wav	269 KB
9.wav	277 KB
10.wav	269 KB
H1 Gab BlueCoin...	
1.wav	241 KB
2.wav	227 KB
3.wav	239 KB
4.wav	293 KB
5.wav	269 KB
6.wav	253 KB
7.wav	321 KB
8.wav	245 KB
9.wav	311 KB
10.wav	287 KB
H1 Gab Headset	
1.wav	253 KB
2.wav	263 KB



True	GussedBaseline	ConfidenceBaselin
"the birch canoe slid on the smooth planks"	"the birth canal we slid on the smooth planks"	0.74933
" a large size in stockings is hard to sell"	"a large size in stockings is hard to sell"	0.9366
"glue the sheet to the dark blue background"	"blue the sheet to the dark blue background"	0.95009
"its easy to tell the depth of a well"	"it's easy to tell the depth of a well"	0.97403
"these days a chicken leg is a rare dish"	"these days a chicken leg is a redfish"	0.86686
"rice is often served in round bowls"	"rice is often served in rum balls"	0.95946
"the juice of lemons makes fine punch"	"the cheese of lemons makes flying punch"	0.79089
"the box was thrown beside the parked truck"	"the box was thrown beside the box truck"	0.82796
"the hogs were fed chopped corn and garbage"	"the Hawks were fed chops corn and garbage"	0.81863
"four hours of steady work faced us"	"4 hours of study world fastest"	0.82847



Plugin Hosting



```
>> noiseRemover = loadAudioPlugin('ERA-N.vst')
```

```
noiseRemover =
```

```
VST plugin 'ERA-N' 2 in, 2 out
```

```
Processing: 40 %
```

```
Gain: 0 dB
```

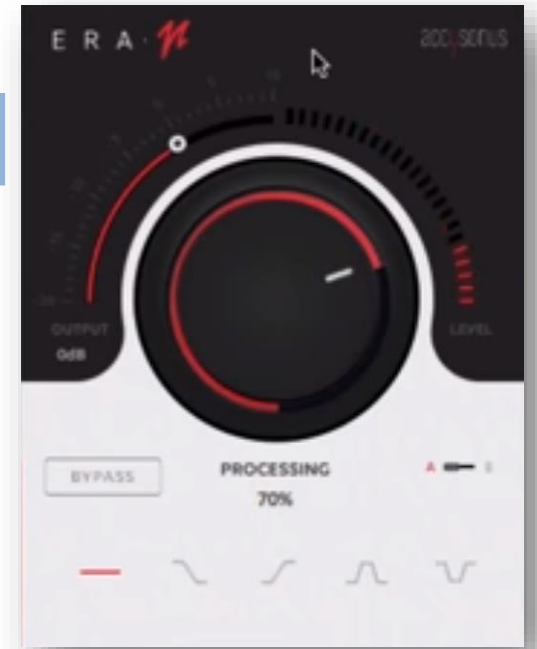
```
Tilt: 'NoTilt'
```

```
Bypass: 0
```

```
>> noiseRemover.Processing = 60;
```

```
>> noiseRemover.Gain = 3;
```

```
>> y = process(noiseRemover, x)
```

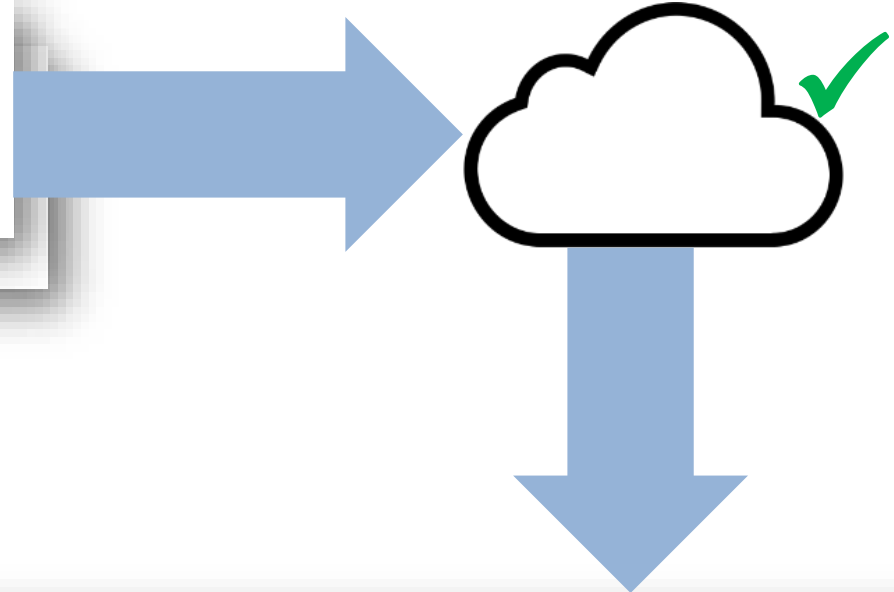
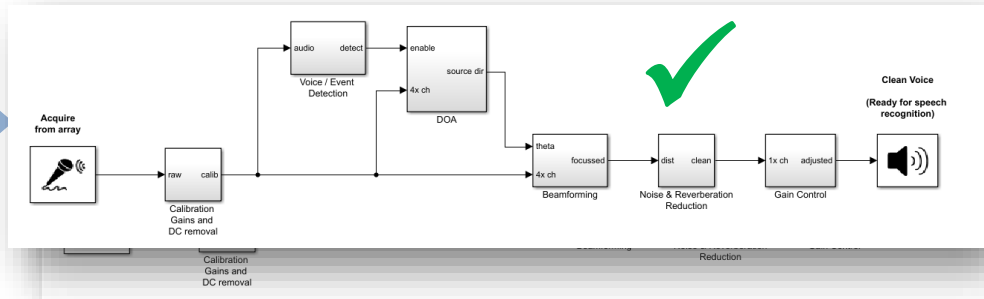


<https://accusonus.com/products/era-n>

Demo

Testing an external audio plugin within a MATLAB live system model

[Video placeholder]



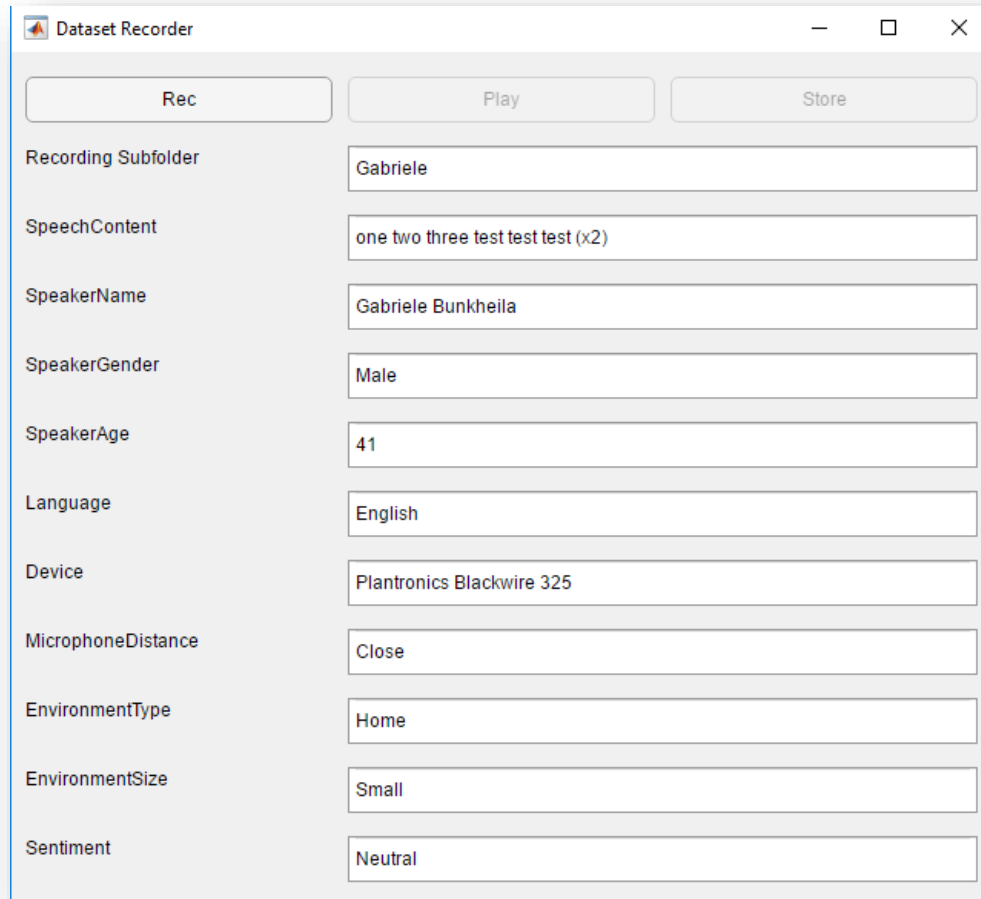
Current Folder

Name	Size
Folder	
Harvard H1	
H1 Gab BlueCoin...	
1.wav	231 KB
2.wav	217 KB
3.wav	209 KB
4.wav	249 KB
5.wav	269 KB
6.wav	251 KB
7.wav	253 KB
8.wav	269 KB
9.wav	277 KB
10.wav	269 KB
H1 Gab BlueCoin...	
1.wav	241 KB
2.wav	227 KB
3.wav	239 KB
4.wav	293 KB
5.wav	269 KB
6.wav	253 KB
7.wav	321 KB
8.wav	245 KB
9.wav	311 KB
10.wav	287 KB
H1 Gab Headset	
1.wav	253 KB
2.wav	263 KB



True	GussedBaseline	ConfidenceBaseline
"the birch canoe slid on the smooth planks"	"the birth canal we slid on the smooth planks"	0.74933
" a large size in stockings is hard to sell"	"a large size in stockings is hard to sell"	0.9366
"glue the sheet to the dark blue background"	"blue the sheet to the dark blue background"	0.95009
"its easy to tell the depth of a well"	"it's easy to tell the depth of a well"	0.97403
"these days a chicken leg is a rare dish"	"these days a chicken leg is a redfish"	0.86686
"rice is often served in round bowls"	"rice is often served in rum balls"	0.95946
"the juice of lemons makes fine punch"	"the cheese of lemons makes flying punch"	0.79089
"the box was thrown beside the parked truck"	"the box was thrown beside the box truck"	0.82796
"the hogs were fed chopped corn and garbage"	"the Hawks were fed chops corn and garbage"	0.81863
"four hours of steady work faced us"	"4 hours of study world fastest"	0.82847

Building a Small Speech Dataset Quickly



Dataset Recorder

Rec Play Store

Recording Subfolder: Gabriele

SpeechContent: one two three test test test (x2)

SpeakerName: Gabriele Bunkheila

SpeakerGender: Male

SpeakerAge: 41

Language: English

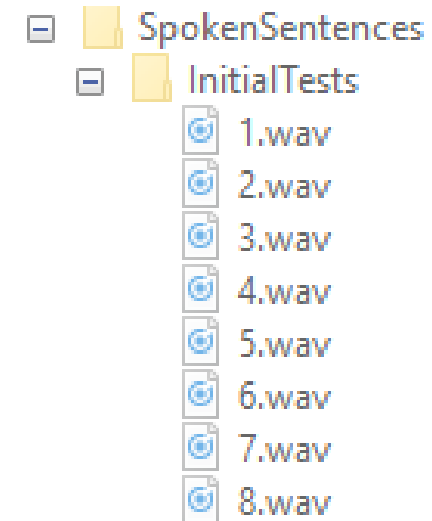
Device: Plantronics Blackwire 325

MicrophoneDistance: Close

EnvironmentType: Home

EnvironmentSize: Small

Sentiment: Neutral



Sentiment	Filename	SampleRate	NumChannels	NumSamples	DateRecorded
"Neutral"	"Neutral\1.wav"	16000	2	73728	11-Sep-2017 11:59:57
"Neutral"	"Neutral\2.wav"	16000	2	62976	11-Sep-2017 12:00:22
"Neutral"	"Neutral\3.wav"	16000	2	1.1674e+05	11-Sep-2017 12:01:00
"Neutral"	"Neutral\4.wav"	16000	2	1.0701e+05	11-Sep-2017 12:01:22
"Neutral"	"Neutral\5.wav"	16000	2	60416	11-Sep-2017 12:01:41
"Neutral"	"Neutral\6.wav"	16000	2	1.9251e+05	11-Sep-2017 12:02:30
"Neutral"	"Neutral\7.wav"	16000	2	1.3107e+05	11-Sep-2017 12:03:04
"Neutral"	"Neutral\8.wav"	16000	2	64512	11-Sep-2017 12:03:23

Building a Small Speech Dataset Quickly

Can you improve speech content labelling?

H1 Harvard Sentences

1. The birch canoe slid on the smooth planks.
2. Glue the sheet to the dark blue background.
3. It's easy to tell the depth of a well.
4. These days a chicken leg is a rare dish.
5. Rice is often served in round bowls.
6. The juice of lemons makes fine punch.
7. The box was thrown beside the parked truck.
8. The hogs were fed chopped corn and garbage.
9. Four hours of steady work faced us.
10. A large size in stockings is hard to sell.

H2 Harvard Sentences

1. The boy was there when the sun rose.
2. A rod is used to catch pink salmon.
3. The source of the huge river is the clear spring.
4. Kick the ball straight and follow through.
5. Help the woman get back to her feet.
6. A pot of tea helps to pass the evening.
7. Smoky fires lack flame and heat.
8. The soft cushion broke the man's fall.
9. The salt breeze came across from the sea.
10. The girl at the booth sold fifty bonds.

H3 Harvard Sentences

1. The small pup gnawed a hole in the sock.
2. The fish twisted and turned on the bent hook.
3. Press the pants and sew a button on the vest.
4. The swan dive was far short of perfect.
5. The beauty of the view stunned the young boy.
6. Two blue fish swam in the tank.
7. Her purse was full of useless trash.
8. The colt reared and threw the tall rider.
9. It snowed, rained, and hailed the same morning.
10. Read verse out loud for pleasure.

H4 Harvard Sentences

1. Hoist the load to your left shoulder.
2. Take the winding path to reach the lake.
3. Note closely the size of the gas tank.
4. Wipe the grease off his dirty face.
5. Mend the coat before you go out.
6. The wrist was badly strained and hung limp.

```
>> harvard{1}
```

```
ans =
```

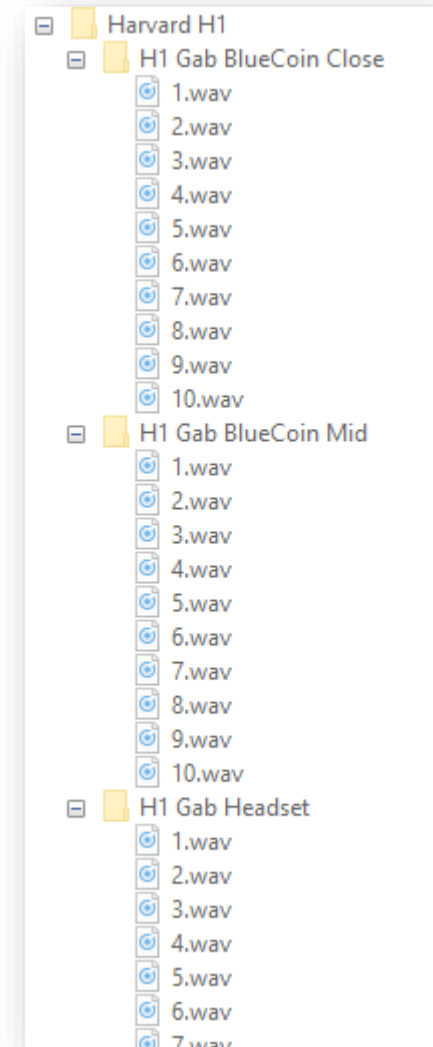
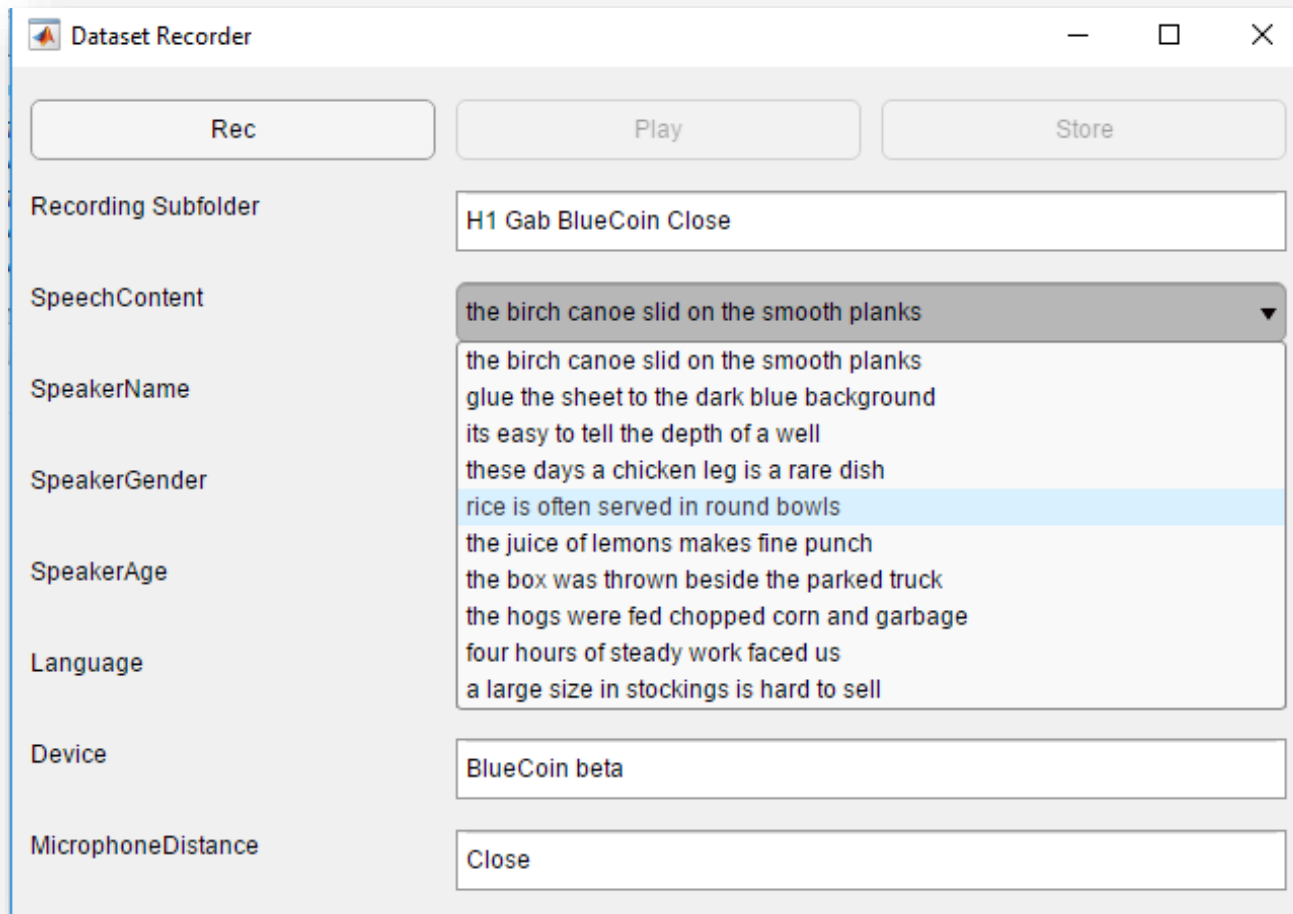
```
10×1 string array
```

```
"the birch canoe slid on the smooth planks"  
"glue the sheet to the dark blue background"  
"its easy to tell the depth of a well"  
"these days a chicken leg is a rare dish"  
"rice is often served in round bowls"  
"the juice of lemons makes fine punch"  
"the box was thrown beside the parked truck"  
"the hogs were fed chopped corn and garbage"  
"four hours of steady work faced us"  
" a large size in stockings is hard to sell"
```

<http://www.cs.cmu.edu/afs/cs.cmu.edu/project/fgdata/OldFiles/Recorder.app/utterances/Type1/harvsents.txt>

Building a Small Speech Dataset Quickly

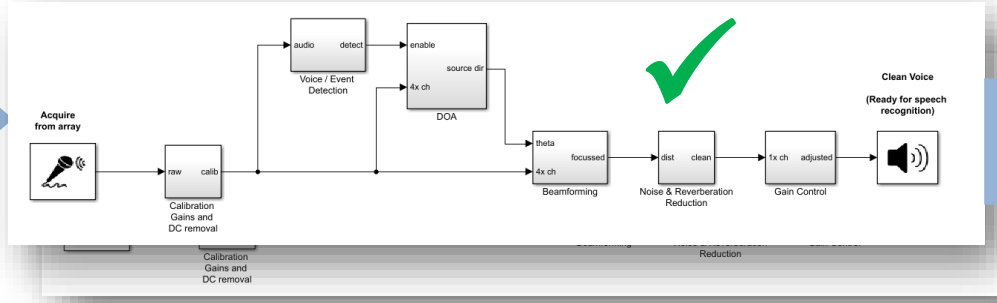
Example: an App with automated content labelling



Demo

Testing speech recognition performance of different design variants (including with new enhancements)

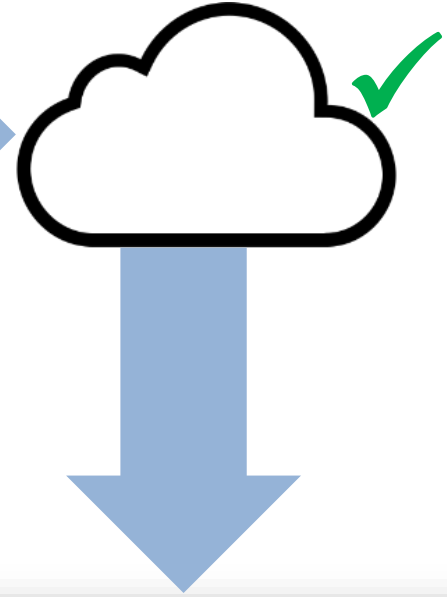
[Video placeholder]



Current Folder

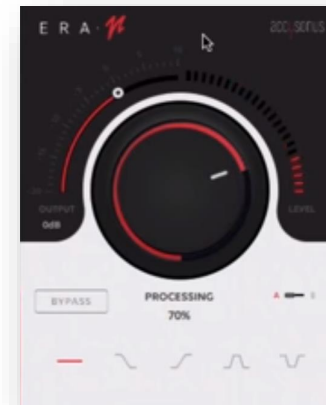
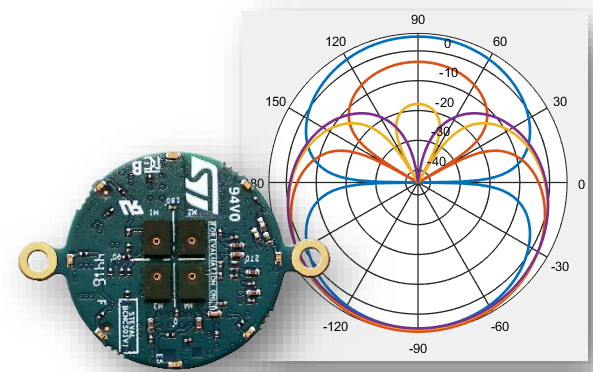
Name	Size
Folder	
Harvard H1	
H1 Gab BlueCoin...	
1.wav	231 KB
2.wav	217 KB
3.wav	209 KB
4.wav	249 KB
5.wav	269 KB
6.wav	251 KB
7.wav	253 KB
8.wav	269 KB
9.wav	277 KB
10.wav	269 KB
H1 Gab BlueCoin...	
1.wav	241 KB
2.wav	227 KB
3.wav	239 KB
4.wav	293 KB
5.wav	269 KB
6.wav	253 KB
7.wav	321 KB
8.wav	245 KB
9.wav	311 KB
10.wav	287 KB
H1 Gab Headset	
1.wav	253 KB
2.wav	263 KB

True	GussedBaseline	ConfidenceBaseline
"the birch canoe slid on the smooth planks"	"the birth canal we slid on the smooth planks"	0.74933
" a large size in stockings is hard to sell"	"a large size in stockings is hard to sell"	0.9366
"glue the sheet to the dark blue background"	"blue the sheet to the dark blue background"	0.95009
"its easy to tell the depth of a well"	"it's easy to tell the depth of a well"	0.97403
"these days a chicken leg is a rare dish"	"these days a chicken leg is a redfish"	0.86686
"rice is often served in round bowls"	"rice is often served in rum balls"	0.95946
"the juice of lemons makes fine punch"	"the cheese of lemons makes flying punch"	0.79089
"the box was thrown beside the parked truck"	"the box was thrown beside the box truck"	0.82796
"the hogs were fed chopped corn and garbage"	"the Hawks were fed chops corn and garbage"	0.81863
"four hours of steady work faced us"	"4 hours of study world fastest"	0.82847



How Can I...

1. Design and simulate a microphone array system?
2. Validate my voice interface can work in real-life scenarios?
3. Understand what else can help me improve my performance?



System Prototype Demo – Summary and Recap

Live voice detection, beamforming and speech recognition

[Video placeholder]

Summary

- Innovate
- Reuse
- Prototype

```

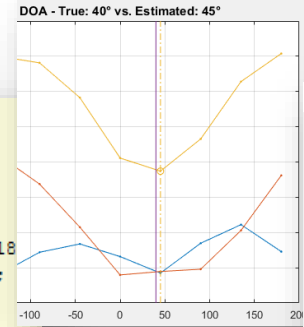
$ Direction to estimate
incidentAngle = [40;0];

$ Theta increments
thetaStep = 45;
steeringThetas = -180:thetaStep:180;
steeringAngles = [steeringThetas;

frameSize = 512;

s = beamformArraysSpecificationsDOA(

```



Dataset Recorder

Rec Play

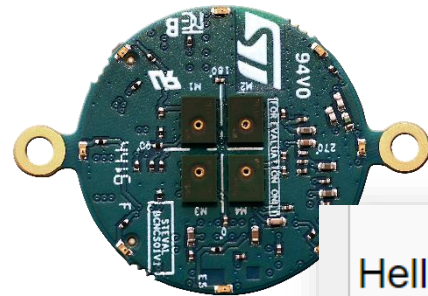
Recording Subfolder: H1 Gab BlueCoin Close

SpeechContent: the birch canoe slid on the

SpeakerName: Gabriele Bunkheila

SpeakerGender: Male

SpeakerAge: 41



Hello
Anyone out there

