



Procedures and analysis methods are described in this guide for the validation of Vesper VA1200 signals in a TWS mockup. Customers can use either the Vesper VA1200 TWS EVK or their own TWS mockups following this guide. Refer to the document “VA1200 Evaluation Kit – TWS Mockup User Guide” for more details on the Vesper VA1200 TWS EVK.

Contents

Test setup	2
Wearing the TWS mockup	2
Simultaneous recording of VA1200 signal and reference microphone signal	2
Test speech and length of data recording.....	2
Calibration tone.....	3
Data recording sequence	3
Data analysis	3
Speech level.....	3
VA1200 SNR.....	4
Test with multiple subjects and TWS mockups with different VA1200 locations	5
VA1200 signal validation criteria	6
What if the VA1200 signal is weak?	7
Appendix	7
Matlab scripts	7
Compute RMS of ref mic signal and SPL of speech	7
Compute SNR of VA1200	7
Gains in signal paths and measurements with physical units	8
Spectral analysis of VA1200 signals	10

Test setup

Wearing the TWS mockup

As shown in the photo, the TWS mockup should be worn snugly in the ear with the proper ear-tip that goes well into the ear canal. If the TWS has a rubber ear-tip, a good check of the fitting of the TWS is the occlusion of the ear – if external sounds are somewhat attenuated, then the earbud is fitting properly. Any flex cable or wires coming from the mockup should be loose (i.e. no tension) and not rubbing against the skin or clothing.



Figure 1 TWS mockup on ear and location of reference microphone

Simultaneous recording of VA1200 signal and reference microphone signal

For VA1200 signal validation, it is required that the VA1200 output and at least one reference microphone output are recorded simultaneously. If there are 2 external microphones on the TWS, use either one.

It is recommended to provide sufficient gain on the VA1200 channel before A/D conversion for the best possible SNR. The Vesper TWS EVK comes with a shoulder board with microphone and VA1200 buffer amplifiers and individually selectable gains. We recommend an amplifier gain of 18 dB on the microphone channels and 42 dB on the VA1200 channel using the Vesper TWS EVK.

The microphone and VA1200 data are normally saved in a multi-channel PCM or WAV file for each recording. For illustrative purpose in the following discussion, we use a 4-channel recording from the Vesper TWS EVK.

Test speech and length of data recording

To cover a wide range of speech frequency and temporal characteristics, we recommend a recording length of approximately 1 minute with half of it containing speech and the other half silence. The speech should not be repeats of a simple phrase such as “1, 2, 3, ...”, rather, it should be

a reading of a set of test phrases, news clips or book paragraphs. The Harvard sentences found here <https://www.cs.columbia.edu/~hgs/audio/harvard.html> are a good choice.

Calibration tone

A 1000 Hz calibration tone is used to calibrate speech level in the recording and is used in recording sessions to prompt the start of voice. Play the tone continuously from a loudspeaker 1 m away from the ear and monitor the sound level with a sound meter placed at the ear. Set the sound meter to A-weighting, slow integration (1 second), then adjust the volume of the tone slowly until the sound meter reads 75 dB SPL. Mark the volume setting and stop the tone.

Data recording sequence

Start data recording -> play the 1 kHz tone at 75 dB SPL for 1 second -> start reading the test phrases -> continue reading for about 30 seconds -> stop reading the test phrases and remain silent and still for about 30 seconds -> stop data recording.

The test voice level should be moderately loud, 70-90 dB SPL at the ear reference point.

Data analysis

An example data recording is shown in Figure 2 where 3 mic channels and 1 VA1200 channel are plotted with labels on different start/stop times. The amplitude has arbitrary units since we may not know the gains in the signal paths.

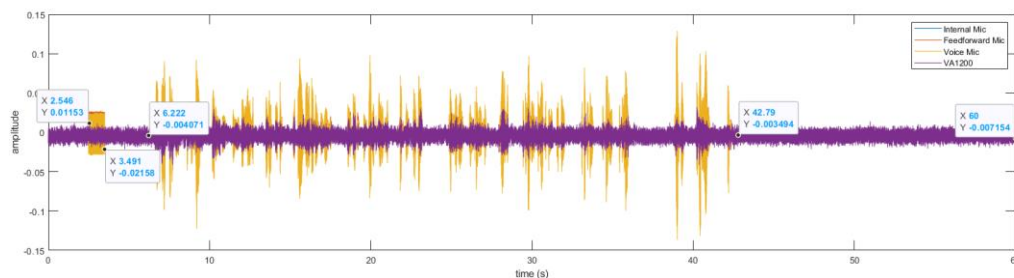


Figure 2 Recording of 4 channels from TWS mockup

Speech level

The 2 external mics, ANC feedforward mic (channel 2) and voice pickup mic (channel 3), have the A4 tone segment (2.546 s – 3.491 s), then the speech segment (6.222 s – 42.790 s) and the silence segment (42.790 s – 60.000 s).

We can compute the speech level from the RMS values of these segments. With channel 2 we have

$$\text{RMS}_{\text{tone}} = 0.01176 \text{ Units}$$

$$\text{RMS}_{\text{speech}} = 0.01116 \text{ Units}$$

and the speech level is

$$20\log(\text{RMS}_{\text{speech}}/\text{RM}_{\text{Stone}}) = 0.45 \text{ dB}$$

over the tone. So, the speech level is:

$$\text{SPL}_{\text{speech}} = (75 + 0.45) \text{ dB SPL} = 75.45 \text{ dB SPL}$$

How to compute the RMS values? In Matlab, a simple script like the one listed in Appendix can be used. In Audacity, you can use Analyze/Contrast shown in Figure 3 to calculate the speech level. Select the tone segment as *Background*, the speech segment as *Foreground*, the *Difference* in dB RMS is the speech level over the tone level. With the values in Figure 3, we have speech level of $75+0.45 = 75.45 \text{ dB SPL}$.

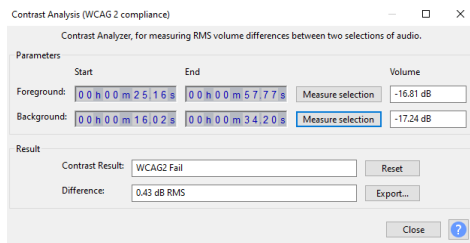


Figure 3 Use Audacity Analyze/Contrast to find speech level

The speech level during the test should be moderately high, at least greater than 70 dB SPL. If the speech level is low, the test subject can try to increase their speech level.

VA1200 SNR

Before computing the RMS values, the VA1200 signal is bandpass filtered. The bandpass filter is normally applied to the VA1200 signal to minimize the VA's self-noise while keeping most of the voice energy, since bone-conducted voice signals picked up by voice accelerometers are band-limited to 100-2400 Hz. This bandpass filtering can be done in Audacity for example. Load the VA1200 signal into Audacity and apply *Effect/High-Pass Filter*, and then *Effect/Low-Pass Filter* to the entire signal, as shown in Figure 4.

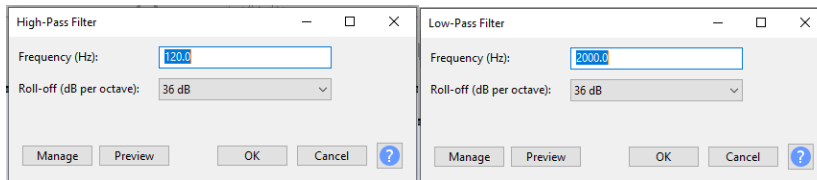


Figure 4 Bandpass filter the VA signal in Audacity using high-pass and low-pass filters

The VA1200 signal in the speech segment (6.222 s – 42.790 s in Figure 2) contains both the bone-conducted voice and the VA's self-noise. The RMS of the noise and RMS of the voice have these relations:

$$(\text{RMS}_{\text{Total}})^2 = (\text{RMS}_{\text{Noise}})^2 + (\text{RMS}_{\text{Voice}})^2 \quad (1)$$

$$\text{RMS}_{\text{Voice}} = \sqrt{(\text{RMS}_{\text{Total}})^2 - (\text{RMS}_{\text{Noise}})^2} \quad (2)$$

$$\text{SNR} = 20\log(\text{RMS}_{\text{Voice}} / \text{RMS}_{\text{Noise}}) + 80 \text{ dB} - \text{SPL}_{\text{speech}} \quad (3)$$

$\text{RMS}_{\text{Total}}$ is the RMS value of the speech segment of the VA1200 data. $\text{RMS}_{\text{Noise}}$ is the RMS value of the silence segment, $\text{RMS}_{\text{Voice}}$ is the value to be computed. As shown in Equation 3, SNR is normalized using the sound pressure level of voice. As the wearer speaks louder, the vibration is higher so the RMS level of the voice section on the VA1200 will be higher relative to that in the silent section. Here, we assume a typical speech level of 80 dB SPL. Therefore, we will increase (or decrease) the SNR by the difference between 80 dB SPL and the measured speech level.

These RMS measurements can be done in Audacity for example. After bandpass filtering the VA1200 signal, open the *Analyze/Contrast* tool (Figure 5), select the speech segment as *Foreground* and silence segment as *Background*. The displayed Difference is 7.64 dB which is the difference between the $\text{RMS}_{\text{Total}}$ and $\text{RMS}_{\text{Noise}}$.

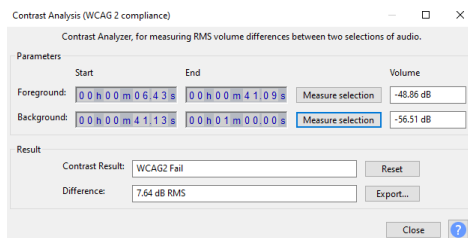


Figure 5 Use Audacity Analyze/Contrast to compute VA1200 SNR

We can use the numbers displayed in the *Contrast Volume* to arrive at true SNR. We first linearize the foreground and background RMS values:

$$\text{RMS}_{\text{Total}} = 10^{-48.86/20}$$

$$\text{RMS}_{\text{Noise}} = 10^{-56.51/20}$$

then use formula (2) and (3) we get the VA1200 signal SNR = 6.83 dB + 80 dB – 75.54 dB = 11.29 dB.

Test with multiple subjects and TWS mockups with different VA1200 locations

We recommend that the TWS mockup is tested with 3 different subjects. Each subject should find the best fitting ear tip and insert the TWS ear tip firmly into his/her ear canal in 3 separate trials. Between trials the TWS mockup should be taken out and re-inserted.

Sometimes, it is desirable to compare VA1200 SNR in two or more different locations. In this case, it is preferred to record the multiple VA1200 channels together with a ref mic channel simultaneously in the same trial if possible.

To handle the full set of subjects and trials, we recommend using an Excel table like the following. The tester can use the methods outlined earlier to fill in the RMS values in the light-shaded cells and the results are calculated automatically. This excel table can also be used to judge if the two locations are statistically the same or different (with 95% confidence). The Excel table is distributed with this guide.

	User 1			User 2			User 3			Average
	Insertion 1	Insertion 2	Insertion 3	Insertion 1	Insertion 2	Insertion 3	Insertion 1	Insertion 2	Insertion 3	
Location #1										
Tone (RMS)	0.01176	0.010584	0.012348	0.0121128	0.0109368	0.009996	0.012936	0.0118776	0.0108192	
Speech (RMS)	0.01116	0.011718	0.0114948	0.0103788	0.009486	0.012276	0.0112716	0.0102672	0.010044	
VA1200 Speech (RMS)	0.0036058	0.0072116	0.0027043	0.0054087	0.0039664	0.0030649	0.0021635	0.0027043	0.0025241	
VA1200 Silent (RMS)	0.0014945	0.0014945	0.0014945	0.0014945	0.0014945	0.0014945	0.0014945	0.0014945	0.0014945	
Calculations:										
Speech SPL	74.5	75.9	74.4	73.7	73.8	76.8	73.8	73.7	74.4	74.5
VA1200 SNR	7.3	12.6	4.2	12.2	9.0	3.3	1.6	4.8	3.3	6.5
Location #2										
Tone (RMS)	0.0119952	0.0107957	0.012595	0.0123551	0.0111555	0.0101959	0.0131947	0.0121152	0.0110356	
Speech (RMS)	0.0109368	0.0114836	0.0112649	0.0101712	0.0092963	0.0120305	0.0110462	0.0100619	0.0098431	
VA1200 Speech (RMS)	0.0050481	0.0100962	0.0037861	0.0075722	0.0055529	0.0042909	0.0035337	0.0037861	0.0035337	
VA1200 Silent (RMS)	0.0014945	0.0014945	0.0014945	0.0014945	0.0014945	0.0014945	0.0014945	0.0014945	0.0014945	
Calculations:										
Speech SPL	74.2	75.5	74.0	73.3	73.4	76.4	73.5	73.4	74.0	74.2
VA1200 SNR	11.0	16.0	8.3	15.6	12.7	7.2	8.2	9.0	7.6	10.6
Is Location #1 Different from Location #2? Yes										
Which Location is Best? #2										
By How Much (dB)? 4.1										

VA1200 signal validation criteria

The table below summarizes the measurements and their limits to validate VA1200 signals. (these numbers should be surveyed from a variety of VA recordings)

Measurement	Limit
VA1200 self-noise (TWS not worn on ear)	~ 1.8 mg (if system gain is used to compute mg)
VA1200 self-noise (TWS worn on ear, no voice)	~ 2.0 mg (if system gain is used to compute mg)
Test speech level at ref mic	70 – 90 dB SPL
VA1200 speech SNR	> 6 dB
Ref mic sound level to VA1200 gain	~ 42 mg / Pa

What if the VA1200 signal is weak?

The weak VA1200 signal in a TWS mockup is mostly a result of poor coupling of the VA1200 to skull vibrations caused by the wearer's own voice production. Several factors can contribute to this poor coupling and the weak signal including:

1. Incorrect size of TWS ear tip.
2. Shallow or loose insertion of the ear tip into the ear canal.
3. Non-optimal location and orientation of the VA1200 in the TWS.
4. Loose mounting of the VA1200 inside the TWS shell such as a floating FPCB.
5. TWS shell material and construction issues; some 3D printed shells are too soft.
6. Low VA1200 signal path gain.
7. The test subject's speech level is too low.
8. System failure.

Remedies to these issues related to the VA1200 location, orientation, and PCB design are described in the Vesper documents "TWS EVK Manual Rev0_0_4" and "VA1200 Location Analysis". Problems related to poor fitting can be remedied by experimenting with difference sizes of ear tips and firmly inserting the ear tips into the ear canal. This guide is intended to be used to evaluate overall system performance and validate signal integrity with the VA1200.

Appendix

Matlab scripts

Compute RMS of ref mic signal and SPL of speech

```
% calculate speech SPL, y is 4-channel recording, ear ref mic is channel 2
toneStart = 2.546; % in seconds, manually determined from graph
toneEnd = 3.491;
speechStart = 6.222;
speechEnd = 42.790;
silenceStart = 42.790;
silenceEnd = 60.000;
RMStone = rms(y(toneStart*Fs:toneEnd*Fs, 2));
RMSspeech = rms(y(speechStart*Fs:speechEnd*Fs, 2));
SPLspeech = 75 + 20*log10(RMSspeech/RMStone);
disp(['Speech level = ' num2str(SPLspeech)]);
```

Notice the rms function is the Signals Toolbox.

Compute SNR of VA1200

```
% calculate VA1200 SNR, channel 4
toneStart = 2.546;
toneEnd = 3.491;
speechStart = 6.222;
speechEnd = 42.790;
silenceStart = 42.790;
silenceEnd = 60.000;
```

```
x = bandpass(y(:,4), [120 2000], Fs, 'ImpulseResponse','iir','Steepness',[0.8 0.8]);
RMSnoise = rms(x(silenceStart*Fs:silenceEnd*Fs));
RMStotal = rms(x(speechStart*Fs:speechEnd*Fs));
RMSvoice = sqrt(RMStotal^2 - RMSnoise^2);
SNRval200 = 20*log10(RMSvoice/RMSnoise);
disp(['VA1200 SNR = ' num2str(SNRval200)]);
```

Notice the rms and bandpass functions are in the Signals Toolbox.

Gains in signal paths and measurements with physical units

The useful goal of the signal measurements is to arrive at physical units of the voice level (dB SPL) and the acceleration (mg) picked up by the VA1200. For this we will need the transducer sensitivities and the total gains of the signal paths. Using the Vesper TWS EVK as an example, we have these parameters:

- VA1200 sensitivity - 45 dBV/g, or scalar sensitivity $S_{VA} = 10^{(-45/20)} \text{ V/g} = 0.0056 \text{ V/g}$
- Reference microphone sensitivity -38 dBV/Pa, or $S_{MIC} = 10^{(-38/20)} \text{ V/Pa} = 0.0126 \text{ V/Pa}$
- VA1200 amplifier gain 42 dB, $A_{VA} = 10^{(42/20)} = 125.89$
- Ref mic amplifier gain 18 dB, $A_{MIC} = 10^{(18/20)} = 7.943$
- A/D is 16-bit, $\pm 5\text{V}$ full scale, hence conversion gain $K_{AD} = (2^{16}\text{bits}) / 10\text{V} = 6553.6 \text{ bits/V}$

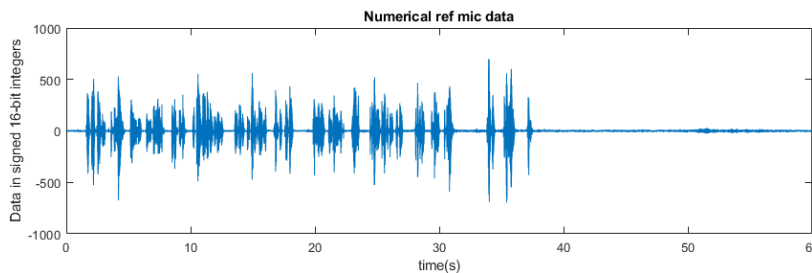


Figure 6 Numerical representation of ref mic data

So, for the ref mic data in 16-bit signed integers, we can calculate the speech in Figure 5 (between 0-40 seconds) RMS value:

- $\text{Speech}_{\text{RMS}} = 69.94 \text{ bits}$, (switch to floating point representation to keep precision)
- Convert the RMS to Pascal at the ref mic:

$$P_{\text{ref}} = \text{Speech}_{\text{RMS}} / K_{AD} / A_{MIC} / S_{MIC} = 0.1067 \text{ Pa}$$

$$\text{which is } 20 \times \log(P_{\text{ref}}/P_0) = 74.54 \text{ dB SPL, where } P_0 = 20 \times 10^{-6} \text{ Pa}$$

VA1200 signal (bandpass filtered to 120 Hz – 2000 Hz) corresponding to the ref mic signal in Figure 2 is shown in Figure 7. Take the silence period (between 40-60 seconds), we have:

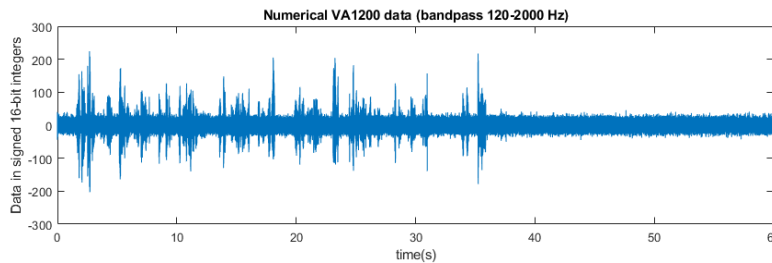


Figure 7 Numerical representation of VA1200 data

- $Noise_{RMS} = 9.677$ bits
- Convert the VA1200 self-noise to mg

$$Acc_{noise} = Noise_{RMS} / K_{AD} / A_{VA} / S_{VA}$$

$$= 0.0021 \text{ g} = 2.1 \text{ mg},$$

which is slightly higher than the noise number in the VA1200 datasheet because there are always vibrations in the real ear.

The VA1200 signal in the speech period (between 0-40 seconds) contains both the bone-conducted voice and the VA's self-noise. It can be shown that the self-noise and the voice signal are uncorrelated and orthogonal, and, with those conditions met, we have:

- $(Total_{RMS})^2 = (Noise_{RMS})^2 + (Voice_{RMS})^2$

$Total_{RMS}$ is just the RMS value of the total VA1200 data between 0-40 seconds, and it is:

- $Total_{RMS} = 22.99$ bits
- And therefore

$$Voice_{RMS} = \sqrt{(Total_{RMS})^2 - (Noise_{RMS})^2} = 20.86 \text{ bits}$$

- Convert the voice pickup to mg

$$Acc_{RMS_{voice}} = Voice_{RMS} / K_{AD} / A_{VA} / S_{VA} = 0.0045 \text{ g} = 4.5 \text{ mg}$$

- And the SNR of VA1200 voice pickup at the test speech level is

$$SNR_{VA} = 20 \times \log(Voice_{RMS} / Noise_{RMS}) = 6.67 \text{ dB}$$

The SNR_{VA} is a critical measurement of the VA1200 signal quality. The other important measurement is the gain from a voice level measured at the ear to the acceleration sensed by the VA1200.

Spectral analysis of VA1200 signals

It is also helpful to examine the signal spectra to validate that the VA1200 is working properly in the TWS mockup. Spectra of the VA1200 noise and voice pickup have distinctive features and can generally fit into the patterns in Figure 8, 9.

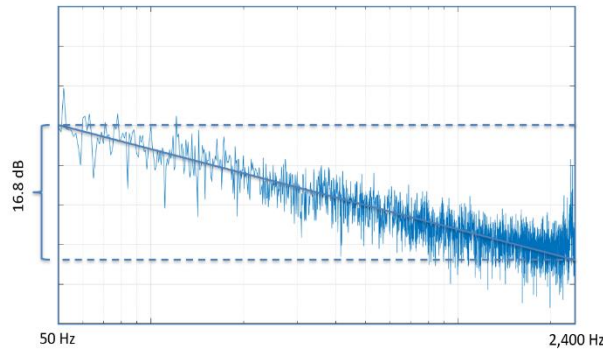


Figure 8 Typical VA1200 self-noise magnitude spectrum. X axis is frequency; Y axis is dB re arbitrary units, assuming channel gain is unknown.

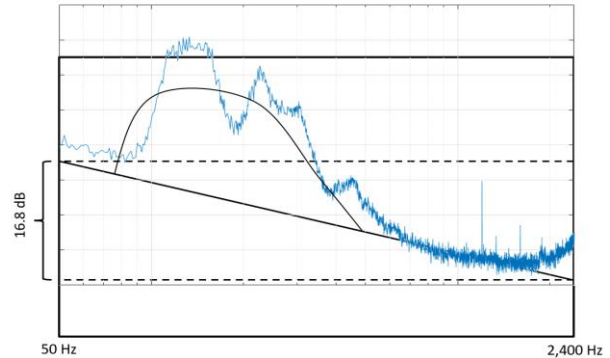


Figure 9 Typical VA1200 voice magnitude spectrum. X axis is frequency; Y axis is dB re arbitrary units, assuming channel gain is unknown.

An example VA1200 recording has the magnitude spectra in Figure 10. The noise spectrum is normally very consistent for frequencies below 1000 Hz. The voice spectrum should be at least 5-6 dB above the noise spectrum for frequencies between 100 and 700 Hz. The general shapes of the spectra help to judge qualitatively whether the VA1200 is working properly, the fitting of the TWS mockup is good and the system is functioning well as a whole.

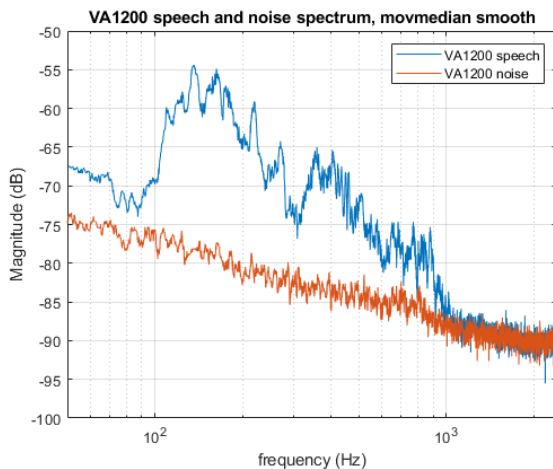


Figure 10 VA1200 noise and voice magnitude spectrum

The Matlab code for the spectrum plot in Figure 10 is the following. Notice the VA1200 signal is not bandpass filtered for this plot and the spectral curves are smoothed with the Matlab function

smoothdata using the moving median filter. The lengths of the time windows of the speech segment and noise segment should be equal.

```
[y Fs] = audioread('demoSubway20210121.wav'); % y is the original 3-channel data, VA1200 on channel 2
x = y(:,2);
speechStart = 22.44;
speechEnd = 34.44;
silenceStart = 33.93;
silenceEnd = 45.93;
[magSpeech, f] = mag(x(speechStart*Fs:speechEnd*Fs), Fs);
% 2 Hz window to smooth curve
wlength = round(2 * (speechEnd - speechStart));
smoothMagSpeech = smoothdata(magSpeech, 'movmedian', wlength);
logSmoothMagSpeech = 20*log10(smoothMagSpeech);
figure;
semilogx(f, logSmoothMagSpeech);
hold on;
[magNoise, f] = mag(x(silenceStart*Fs:silenceEnd*Fs), Fs);
wlength = round(2 * (silenceEnd - silenceStart));
smoothMagNoise = smoothdata(magNoise, 'movmedian', wlength);
logSmoothMagNoise = 20*log10(smoothMagNoise);
semilogx(f, logSmoothMagNoise);
hold off;
title('VA1200 speech and noise spectrum, movmedian smooth');
xlim([50 2400]);
grid on;
xlabel('frequency (Hz)');
ylabel('Magnitude (dB)');
legend('VA1200 speech', 'VA1200 noise', 'Location', 'NorthEast');

function [m, f] = mag(x, Fs)
n = length(x);
if rem(n,2)==1
    x = x(1:end-1);
    n = n - 1;
end
X = fft(x);
k = n/2; % scaling factor
f = (0:n-1)*(Fs/n);
f = f(1:n/2); % only half fft is needed
m = abs(X(1:n/2))/k;
m(1) = m(1)/2; % half the zero freq point
end
```