

DSZOB, cvičenie 7.

Zadanie:

Úloha 1 - Formantová analýza

Načítajte si nahrávku z dokumentového servera (vowels.m4a).

Vykonajte analýzu jednotlivých samohlások pomocou spektrogramu v danej nahrávke. Identifikujte hodnoty jednotlivých formantov. Na základe tabuľky sa pokúste identifikovať o aké anglické samohlásky ide. Overte si výpočet výpočtom nahrávky.

Ako pomôcku môžete použiť nasledujúcu tabuľku:

Short vowels	/a/		/e/		/i/		/o/		/u/	
	Mean	sd	Mean	sd	Mean	sd	Mean	sd	Mean	sd
F1	548	80	448	44	367	48	441	34	355	32
F2	1268	132	1899	194	2086	73	1009	130	1370	313
F3	2519	395	2985	583	3215	449	2659	394	2556	368
Long vowels	/a:/		/e:/		/i:/		/o:/		/u:/	
	Mean	sd	Mean	sd	Mean	sd	Mean	sd	Mean	sd
F1	666	70	448	38	326	26	451	45	345	21
F2	1156	49	2048	95	2178	74	860	152	1213	223
F3	2287	266	2946	449	2971	272	2402	362	2330	165

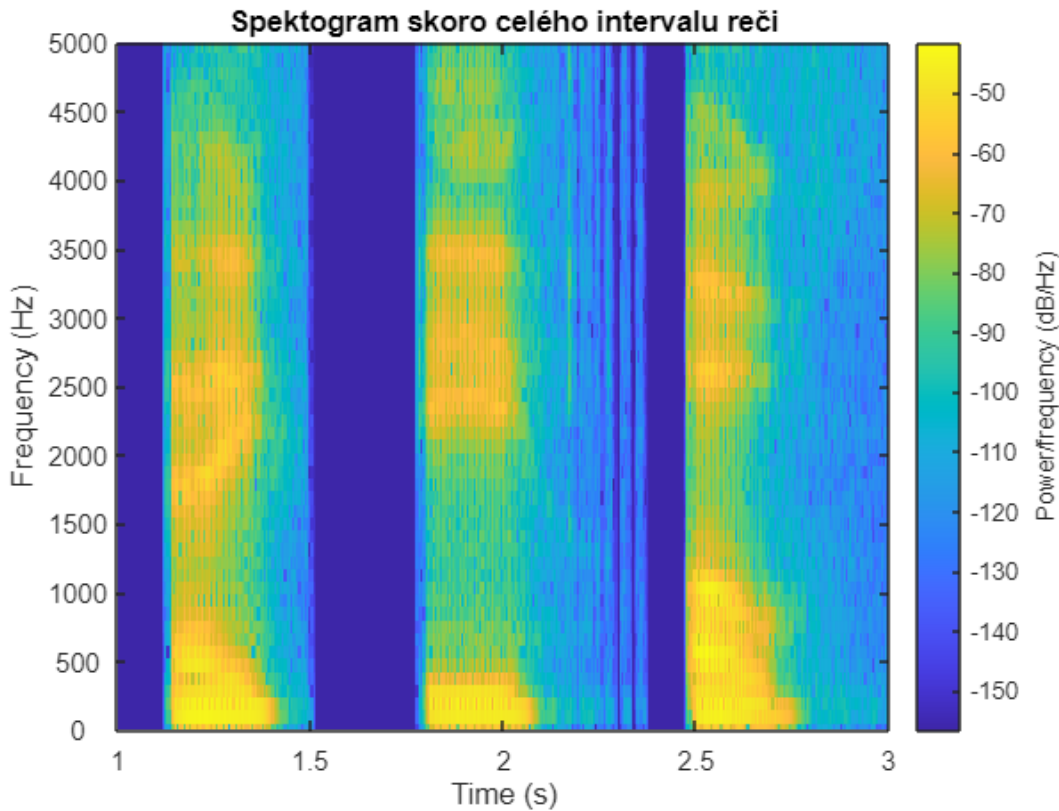
Solution:

1.1 Načítali sme si nahrávku z dokumentového servera (vowels.m4a).

```
% Riešenie / Solution
speech = importdata("vowels.m4a");
% speech = importdata("Hlas1.m4a");
fs = speech.fs;
rawSignal = speech.data();
rawSignal = rawSignal';
rawSignalLength = length(rawSignal);
maxTime = rawSignalLength / fs;
timeline = 0:maxTime/(rawSignalLength-1):maxTime;
```

1.1.1 Ktorej sme vizualizovali spektrálny graf, v úseku kde sa nachádzajú prvé tri úseky hlások.

```
figure
spectrogram(rawSignal, hamming(512), 256, 512, fs, "yaxis")
title('Spektrogram skoro celého intervalu reči')
set(gca, 'YTickLabel', (0:5000/10:5000))
ylabel('Frequency (Hz)')
ylim([0 5])
xlim([1 3])
```



1.1.2 Tieto úseky hlásom sme ešte raz vizualizovali, ale osobitne

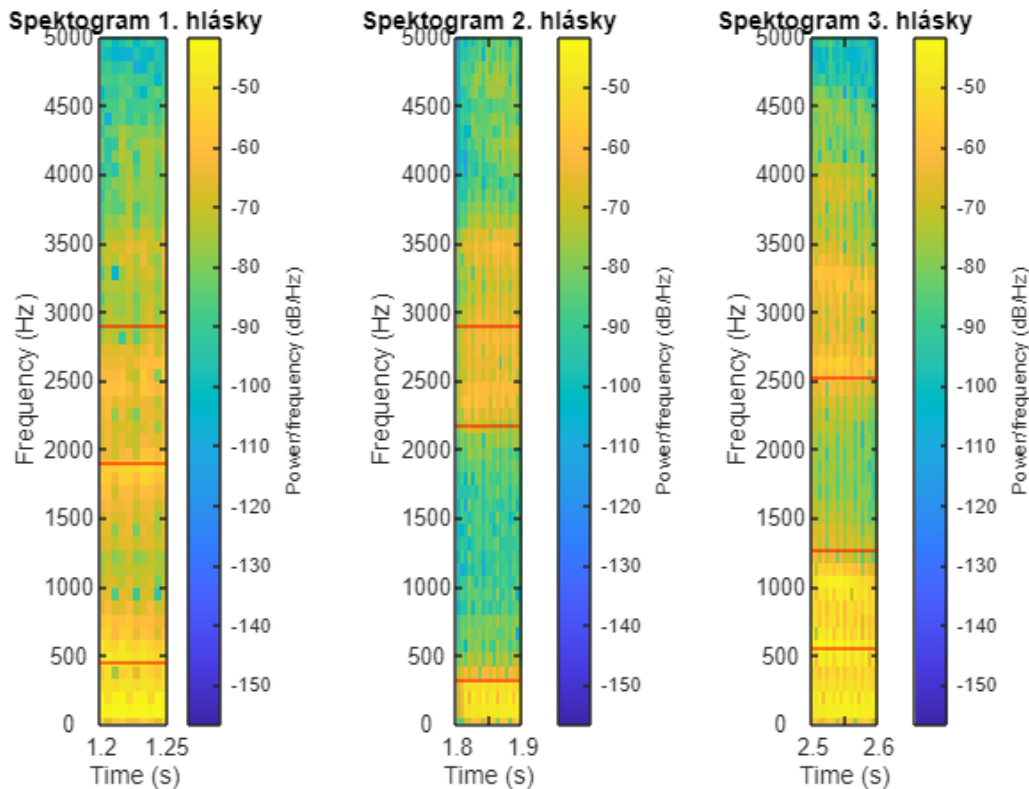
```
figure
    tiledlayout(1,3)
        nexttile
        spectrogram(rawSignal, hamming(512), 256, 512, fs, "yaxis")
        title('Spektrogram 1. hlásky')
        set(gca, 'YTickLabel', (0:5000/10:5000))
        ylabel('Frequency (Hz)')
        ylim([0 5])
        xlim([1.2 1.25])
        hold on
        yline([0.448 1.899 2.895], Color="red")
        hold off

        nexttile
        spectrogram(rawSignal, hamming(512), 256, 512, fs, "yaxis")
        title('Spektrogram 2. hlásky')
        set(gca, 'YTickLabel', (0:5000/10:5000))
        ylabel('Frequency (Hz)')
        ylim([0 5])
        xlim([1.8 1.9])
        hold on
        yline([0.326 2.178 2.9], Color="red")
        hold off
```

```

nexttile
spectrogram(rawSignal, hamming(512), 256, 512, fs, "yaxis")
title('Spektrogram 3. hlásky')
set(gca,'YTickLabel', (0:5000/10:5000))
ylabel('Frequency (Hz)')
ylim([0 5])
xlim([2.5 2.6])
hold on
yline([0.548 1.268 2.519],Color="red")
hold off

```



```

rec1 = audioplayer(rawSignal, fs);

playblocking(rec1, [round(1*fs),round(1.5*fs)]) %ej
playblocking(rec1, [round(1.8*fs),round(2.3*fs)]) %i
playblocking(rec1, [round(2.4*fs),round(2.8*fs)]) %ou

```

Nájdené **Formanty** u jednotlivých hlásky:

1. cca. 500Hz, 1800-2000Hz, 2480-2500Hz
2. cca. 300-400HZ, 2200-2500Hz, 2700-2900hz,
3. cca. 500Hz, 1000Hz, 2600Hz

Podľa týchto hŕdnutí odhadujem písmena pre jednotlivé hlásky nasledovne:

1. Podľa podmienok toho že tam nie sú blízke hodnoty okolo 1000, ale tesno pri 3000, tak vypadáva **a o u**, a najpravdepodobnejšie hláska je **short e**, vzhľadom na F2 blízke 2000 ale stále pod ním, **s F1=448, F2=1899, F3=2985**, až na ten posledný parameter
2. Opäť vyradovacou metódou to podľa podmienok nebude **a o** ani **u**, je to pravdepodobne **long i**, s hodnotami **F1=326, F2=2178, F3=2900**
3. Podľa prvých troch približných hodnôt, vypadáva **e i**, pričom najpravdepodobnejšia hláska bola identifikovaná táto hláska ako **short a**, s **F1=548, F2=1268, F3=2519**, alebo **o**

Podľa nahrávok určujem tieto hlásky nasledovne:

1. **ej**
2. **i**
3. **ou**

Je potrebné poznamenať že síce hlásky by mali byť anglické, zápis /a/ prislúcha k vyslovnosti, teda blood sa vyslovuje bl/a/d, a /i:/ keep, k/i:/p. Vzhľadom na to boli porovnávané samohlásky ako znejú v slovenčine a nie ich anglická varianta.

Vyhodnotenie:

- Identifikované hlásky boli iba približne polovične zhodné s realne vypočítanými hláskami, to bolo jednak tým že v jednom úseku hlásky boli konkrétne 2 hlásky, tj. pri 1. a 3. úseku hlásky a tým že pri analýze 3. úseku hlásky, boli určené formanty podobné v 2 hláskach.

Zadanie 2 - Fundamentálna frekvencia

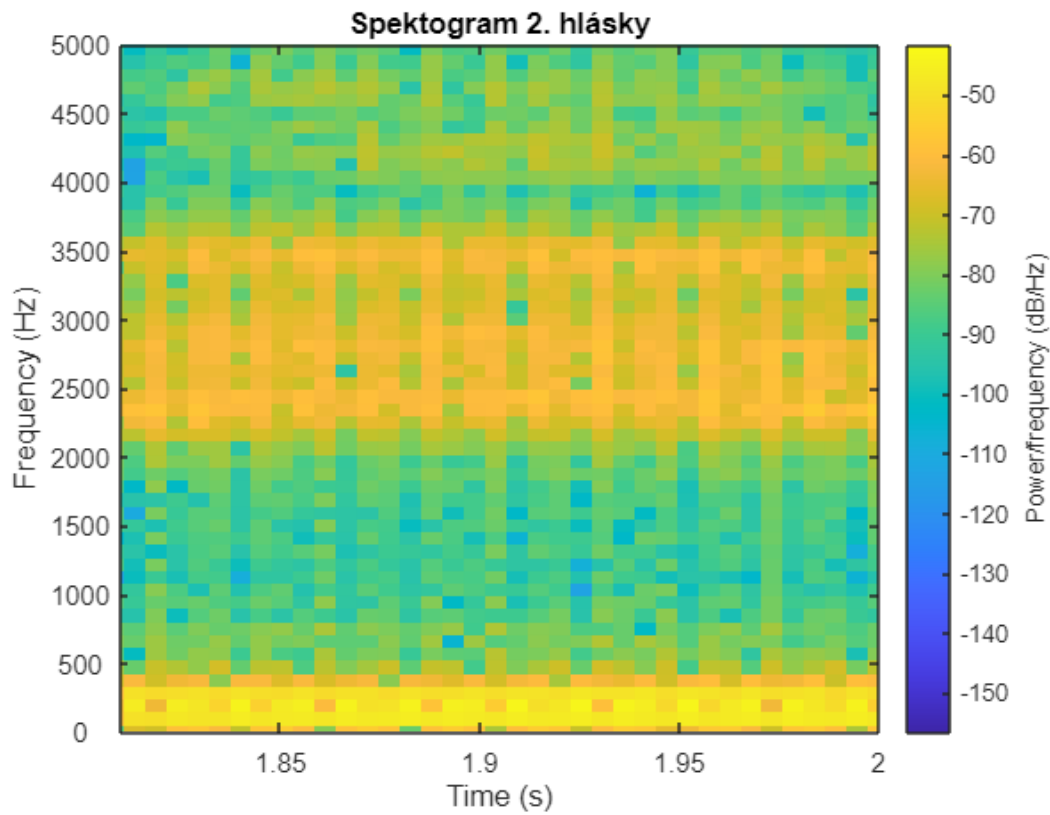
Pre jednu z hlások z nahrávky vykonajte odhad fundamentálnej frekvencie pomocou complexného cepstra.

1. Vezmite okienko 2048 hodnôt zo stredu zvolenej samohlásky
2. Pre toto vybrané okno vypočítajte complexné cepstrum (cceps)
3. Vizualizujte peak vo výsledku v rozsahu približne 0.002 - 0.01s
4. Vypočítte príslušnú frekvenciu nájdeného peaku ($f = 1/T$)

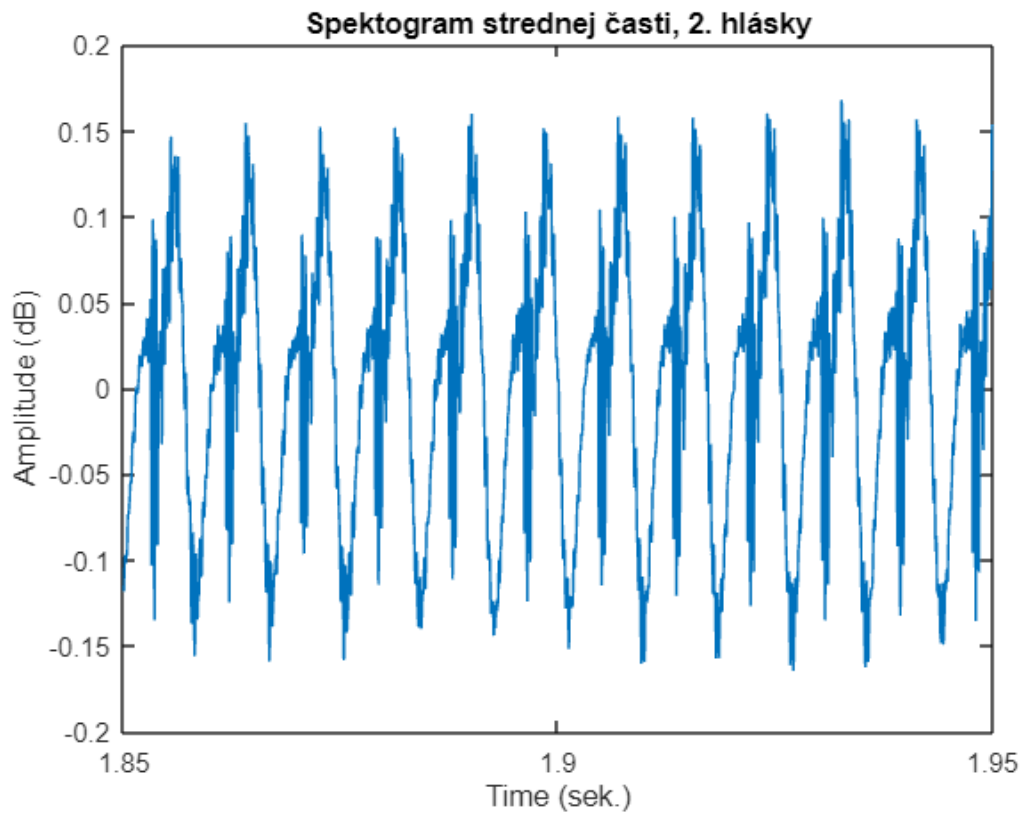
% Riešenie / Solution

2:1 Vzali sme si okienko s 2048 hodnôtami zo stredu zvolenej samohlásky -->

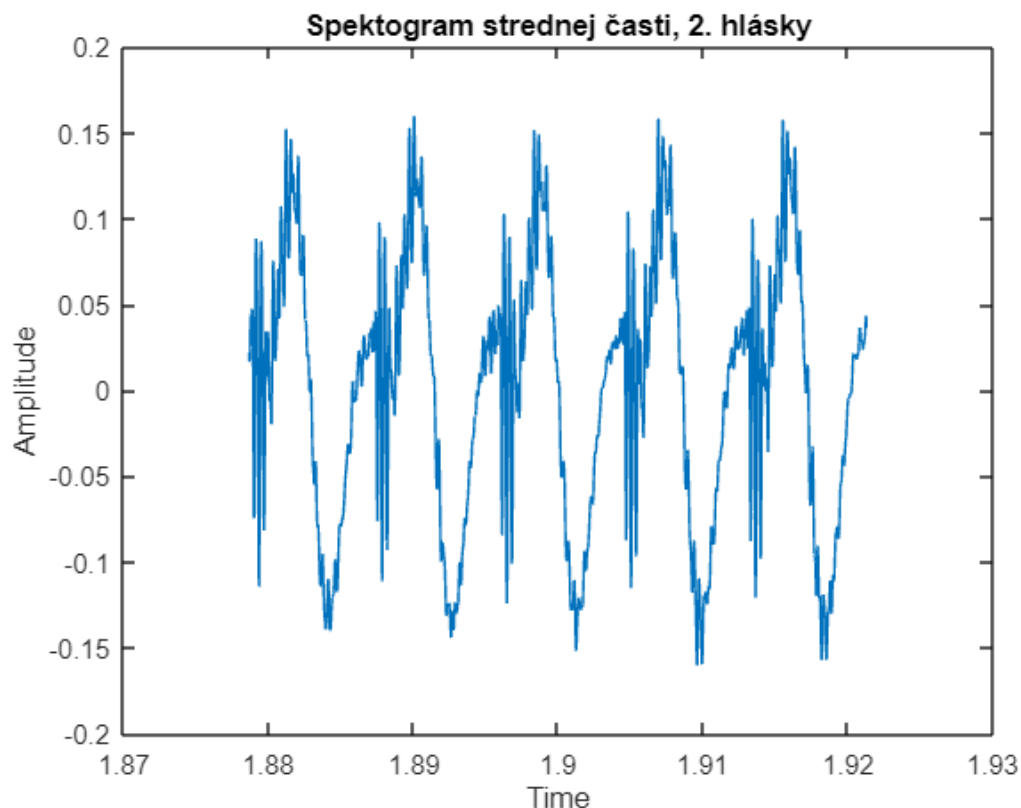
```
% Výber úseku 2. hlásky
figure
spectrogram(rawSignal, hamming(512), 256, 512, fs, "yaxis")
title('Spektrogram 2. hlásky')
set(gca, 'YTickLabel', (0:5000/10:5000))
ylabel('Frequency (Hz)')
ylim([0 5])
xlim([1.81 2])
```



```
% priblíženie úseku 2. hlásky
figure
    chosenIndx = 1.85*fs+1:1.95*fs;
    chosenInterval = rawSignal(chosenIndx);
    chosenTimeline = timeline(chosenIndx);
    plot(chosenTimeline,chosenInterval)
    title('Spektrogram strednej časti, 2. hlásky')
    xlabel("Time (sek.)")
    ylabel("Amplitude (dB)")
```



```
% výber okna dĺžky 2048 zo strednej časti, vybraného úseku 2. hlásky
halfOfChosenSignal = round(length(chosenInterval) / 2);
chosenIndx = halfOfChosenSignal - 1023:halfOfChosenSignal + 1024;
chosenTimeline = chosenTimeline(chosenIndx);
figure
    chosenInterval = chosenInterval(:,chosenIndx);
    plot(chosenTimeline,chosenInterval)
    title('Spektrogram strednej časti, 2. hlásky dĺžky 2048 time samples')
    xlabel("Time (sek.)")
    ylabel("Amplitude")
```

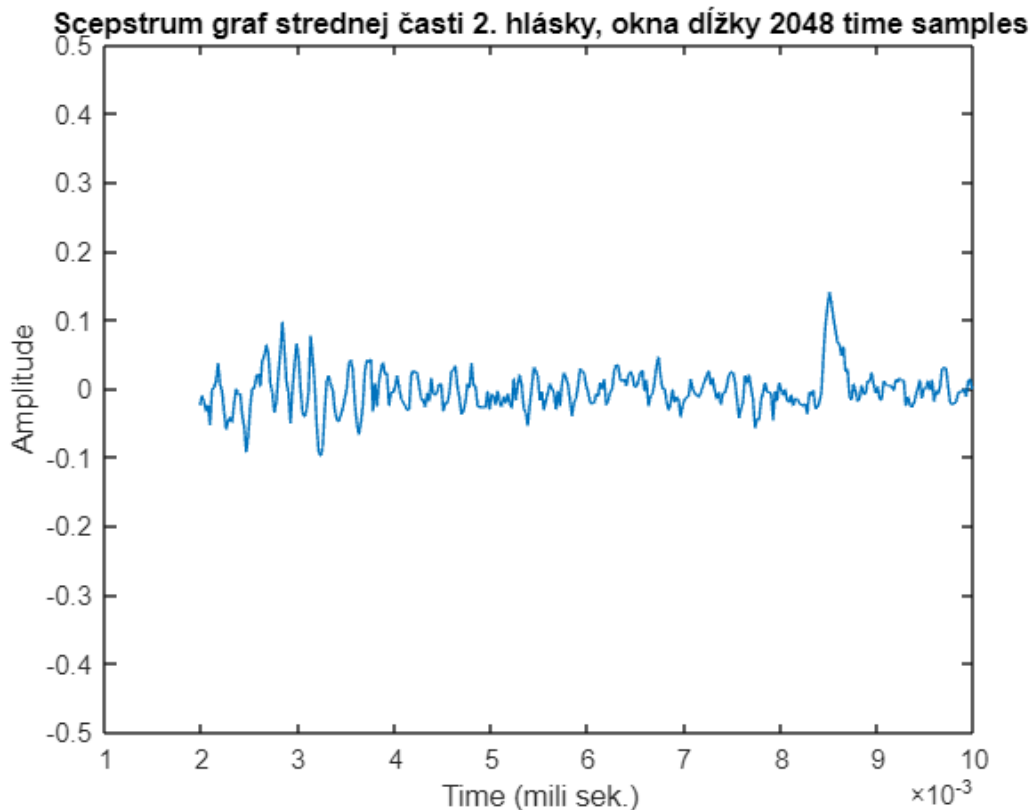


2.2 Pre toto vybrané okno sme vypočítali complexné cepstrum (pomocou funkcie cceps, do ktorej sme vložili ako jediný parameter, vybrané okno 2048 2. hlásky)

```
% Výpočet cepstrum za účelom získania F0, to je získané po nájdení
% lokálneho maxima pre interval 0.002s až 0.01s výsledného cepstra
cepstrum = cceps(choosenInterval);
approximateInterval = 0.002*fs:0.01*fs;
```

2.3 Vizualizovali sme graf v rozsahu približne 0.002 - 0.01s výsledného cepstrum vektora, kde sme identifikovali peak

```
figure
    plot(choosenTimeline(approximateInterval) -
    choosenTimeline(1),cepstrum(approximateInterval))
    ylim([-0.5 0.5])
    title('Scepstrum graf strednej časti 2. hlásky, okna dĺžky 2048 time samples')
    xlabel("Time (mili sek.)")
    ylabel("Amplitude")
```



2.4 Vypočítali sme príslušnú frekvenciu nájdeného peaku pomocou vzorca ($f = 1/T$)

```
% Nájde si dané maximum a jeho index
[M, mIndx] = max(cepstrum(approximateInterval))
```

```
M = 0.1397
mIndx = 314
```

```
tempInt = chosenTimeline(approximateInterval);
T = (tempInt(mIndx)-tempInt(1))
```

```
T = 0.0065
```

```
% a použijeme ich pre výpočet frekvencie pre nájdený index, pomocou
% vzorca  $f = 1/T$ 
% Kedy periodu sme určili vrámci 2. hlásky, preto sme odčítali čas kedy
% sa začala
freq = 1/T;
% Takto sme získali frekvenčnú hodnotu  $F0 = 153\text{Hz}$ 
display("F0 = " + freq + "Hz")
```

```
"F0 = 153.3537Hz"
```