Jaroslav Reichl, © 2011

ANALÝZA LIDSKÉHO HLASU

Pomůcky

mikrofon MCA-BTA, LabQuest, program LoggerPro (nebo LoggerLite), tabulkový editor Excel, program Mathematica

Postup

Z každodenní zkušenosti víme, že každý lidský hlas je jiný a přesto mají jednotlivé hlásky, které konkrétní člověk vydává, jakýsi společný základ. Co je tím základem a jak dopadne porovnání hlasu dospělého muže a mladé dívky, jsem se pokoušel zjistit.

Postupně jsem pomocí mikrofonu firmy Vernier zaznamenal průběh závislosti intenzity zvuku na čase pro všechny samohlásky, tj. *a*, *e*, *i*, *o* a *u*. Mikrofon jsem připojil k LabQuestu, ponechal standardní nastavení vzorkovací frekvence a doby měření a vyslovil jsem příslušnou samohlásku. Zvuk samohlásky jsem vyslovoval spojitě po dobu několika sekund, během nichž jsem spustil měření na LabQuestu. Relativně dlouhé trvání vyslovování dané hlásky vyplývalo z toho, že LabQuest potřebuje určitou dobu na to, aby zahájil měření. Po úspěšném záznamu dané hlásky jsem naměřená data uložil a provedl další měření.

Analogicky jsem pak postupoval s nahráváním samohlásek Šárky Lorencové, žákyně SPŠST Panská v Praze. Naměřená data jsem potom importoval do programu LaggerPro a odtud dále do tabulkového editoru Excel. Tento import byl nutný pouze jako mezikrok, neboť jsem chtěl data zpracovávat v programu Mathematica a ten umí načítat data např. právě z editoru Excel; z programu LoggerPro je načíst do programu Mathematica nelze.

Zobrazit v programu Mathematica naměřená data bylo snadné. Komplikovanější bylo zjistit frekvence harmonických složek, ze kterých lze zvuk samohlásek složit. Pro okamžitou výchylku *y* v čase *t* naměřeného signálu můžeme totiž psát

$$y = b_0 + \sum_{\mathrm{i}=1}^{\infty} \left(a_\mathrm{i} \sin\left(2\pi f_\mathrm{i} t\right) + b_\mathrm{i} \cos\left(2\pi f_\mathrm{i} t\right) \right), \tag{1}$$

kde koeficienty a_i a b_i pro $i = 1, 2, \ldots$ udávají váhu, s jakou jsou jednotlivé harmonické složky ve zvuku dané samohlásky zastoupeny. Koeficient b_0 představuje pouze posun po ose *y* a je proto pro frekvenční analýzu nezajímavý.

Abychom získali frekvence jednotlivých harmonických složek obsažených ve výrazu (1), bylo nutné provést Fourierovu analýzu daného signálu. Tento signál byl ovšem nespojitý, neboť LabQuest zaznamenával data s určitou vzorkovací frekvencí (v tomto případě s frekvencí $f_\mathrm{vz} = 10\ \mathrm{kHz}$). Proto bylo nutné provést tzv. diskrétní Fourierovu analýzu.

Označíme-li $y(t_\mathrm{n})$ okamžitou výchylku zaznamenávaného signálu v čase t_n a $A(f_\mathrm{k})$ amplitudu frekvence f_k ve frekvenčním spektru uvažovaného signálu, můžeme pro amplitudu $A(f_\mathrm{k})$ na základě vztahů pro diskrétní Fourierovu transformaci psát

$$A(f_\mathrm{k}) = T \sum_{\mathrm{n}=0}^{\mathrm{N}-1} x(t_\mathrm{n}) e^{-\frac{2\pi}{\mathrm{N}} \mathrm{n} \cdot \mathrm{k} \cdot \mathrm{i}}, \tag{2}$$

kde *N* je počet vzorků signálu, které analyzujeme, $T = \dfrac{1}{f_\mathrm{vz}}$ je perioda vzorkování signálu, i je imaginární jednotka s vlastností $\mathrm{i}^2 = -1$ a $k = 0, 1, 2, \ldots, N-1$. V případě diskrétní Fourierovy analýzy se vyplatí místo funkcí sinus a kosinus vyjádřit harmonické složky pomocí exponenciální funkce s komplexním argumentem na základě Eulerova vztahu

$$e^{\mathrm{i}\cdot\varphi} = \cos\varphi + \mathrm{i}\sin\varphi\ ; \tag{3}$$

v tom případě amplitudy $A(f_\mathrm{k})$ použité ve vztazích (2), (4), (5) a (6) mají význam Fourierových koeficientů a_i a b_i (pro $i = 0, 1, 2, ..., N-1$) použitých ve vztahu (1).

Pro zpětnou (inverzní) Fourierovu transformaci pak můžeme psát

$$y(t_\mathrm{n}) = \frac{1}{NT}\sum_{\mathrm{k}=0}^{\mathrm{N}-1} A(f_\mathrm{k})\, e^{\frac{2\pi}{\mathrm{N}}\mathrm{n}\cdot\mathrm{k}\cdot\mathrm{i}}\,, \tag{4}$$

kde $n = 0, 1, 2, ..., N-1$.

Při výpočtech se v praxi používají místo vztahů (2) a (4) zjednodušené vztahy, které jsou vhodnější např. i pro počítačové zpracování naměřeného signálu. Místo vztahu (2) tedy můžeme psát

$$A(f_\mathrm{k}) = \sum_{\mathrm{n}=0}^{\mathrm{N}-1} x(t_\mathrm{n})\, e^{-\frac{2\pi}{\mathrm{N}}\mathrm{n}\cdot\mathrm{k}\cdot\mathrm{i}} \tag{5}$$

a místo vztahu (4) můžeme psát vztah

$$y(t_\mathrm{n}) = \frac{1}{N}\sum_{\mathrm{k}=0}^{\mathrm{N}-1} A(f_\mathrm{k})\, e^{\frac{2\pi}{\mathrm{N}}\mathrm{n}\cdot\mathrm{k}\cdot\mathrm{i}}\,. \tag{6}$$

Pro samotnou frekvenční analýzu je důležitý vztah (5). Je důležité si ovšem uvědomit, že vypočtené hodnoty $A(f_\mathrm{k})$ jsou komplexní čísla. Proto je nutné při zobrazování frekvenčního spektra zobrazovat na svislou osu absolutní hodnotu z dané hodnoty $A(f_\mathrm{k})$. Vztah (6) slouží k ověření, že amplitudy jednotlivých frekvencí vypočtené na základě vztahu (5) jsou správné. Aplikací vztahu (6) na právě vypočtené amplitudy frekvencí musíme dostat zpět původní naměřený signál. To je také v připraveném notebooku [5] programu Mathematica ověřeno a u všech zkoumaných zvuků je to splněno.

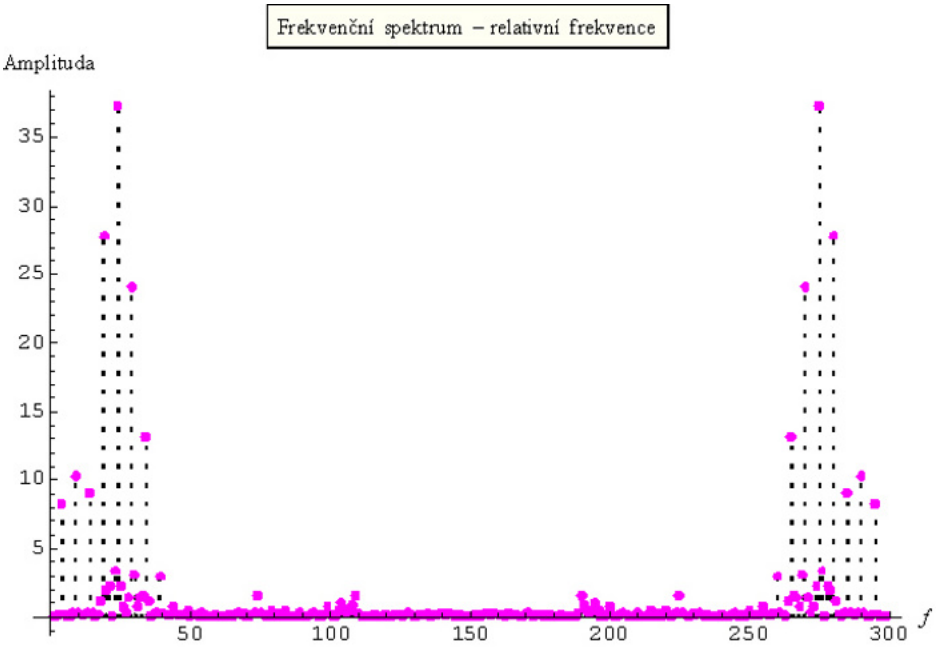
Pomocí vztahu (5) tedy získáme amplitudy harmonických frekvencí, které se v naměřeném signálu vyskytují. Problémem je, že v diskrétní Fourierově transformaci mají frekvence celočíselné hodnoty a nevyjadřují tedy skutečné frekvence harmonických složek, které jsou obsaženy v signálu popsaném vztahem (1). Proto je nutné tyto hodnoty převést na skutečné frekvence $f_{\mathrm{k\ skutečná}}$ vztahem

$$f_{\mathrm{k\ skutečná}} = \frac{k}{t_{\mathrm{celkový}}}\,, \tag{7}$$

kde $k = 1, 2, ..., N$ odpovídá pořadí frekvence f_k, jejíž amplituda $A(f_\mathrm{k})$ je určená vztahem (5), a $t_{\mathrm{celkový}}$ je celkový čas, po který měření probíhalo. Pro frekvenční analýzu mají význam pouze skutečné frekvence, které jsou definované vztahem (7), pro $k = 2, 3, ..., N$. Pro $k = 1$ totiž získáváme absolutní člen výrazu (1), který nemá pro frekvenční analýzu význam.

(V notebooku [5] programu Mathematica se pracuje z praktických důvodů s proměnnými *k* a *n* v rozsahu $1, 2, ..., N$, ačkoliv ve vztazích (2), (4), (5) a (6) nabývají tyto proměnné hodnot $0, 1, 2, ..., N-1$.)

Při základním zpracování Fourierovy analýzy získáme frekvenční spektrum, které je symetrické (viz obr. 1) a to díky použité metodě výpočtu pomocí vztahu (5). Z fyzikálního hlediska má smysl pouze první část spektra. Proto jsem v programu Mathematica připravil možnost zobrazit detailně pouze výřez frekvenčního spektra, ve kterém budou zobrazeny pouze fyzikálně zajímavé frekvence harmonických složek. Současně s tím budou vypsány hodnoty těch frekvencí, které se na barvě dané samohlásky podílejí nejvíce. To je v notebooku [5] programu Mathematica řešeno tak, že nastavíme na určitou hodnotu, kterou odečteme z grafu analogického tomu na obr. 1, proměnnou *hladinasumu*. V případě grafu zobrazeného na obr. 1 bychom volili nastavení *hladinasumu* = 5 a *pocetvybranychdat* = 100.



obr. 1

Zdroje

[1] FEYNMAN R., LEIGHTON R. B., SANDS M. *Feynmanove prednášky z fyziky*, Alfa, vydavateľstvo technickej a ekonomickej literatúry Bratislava, 1985

[2] HLAVÁČ V., SEDLÁČEK M. *Zpracování signálů a obrazů*, nakladatelství ČVUT Praha, 2007

[3] AUERSVALD A., BALIHAR T. *Hudební akustika*, maturitní práce, SPŠST Panská, Praha, 2011

[4] Základy Fourierovy analýzy v Encyklopedii fyziky

[5] videonávod na export dat z programu LoggerPro do programu Mathematica

[6] zdrojová data se záznamem měření se senzorem GPS

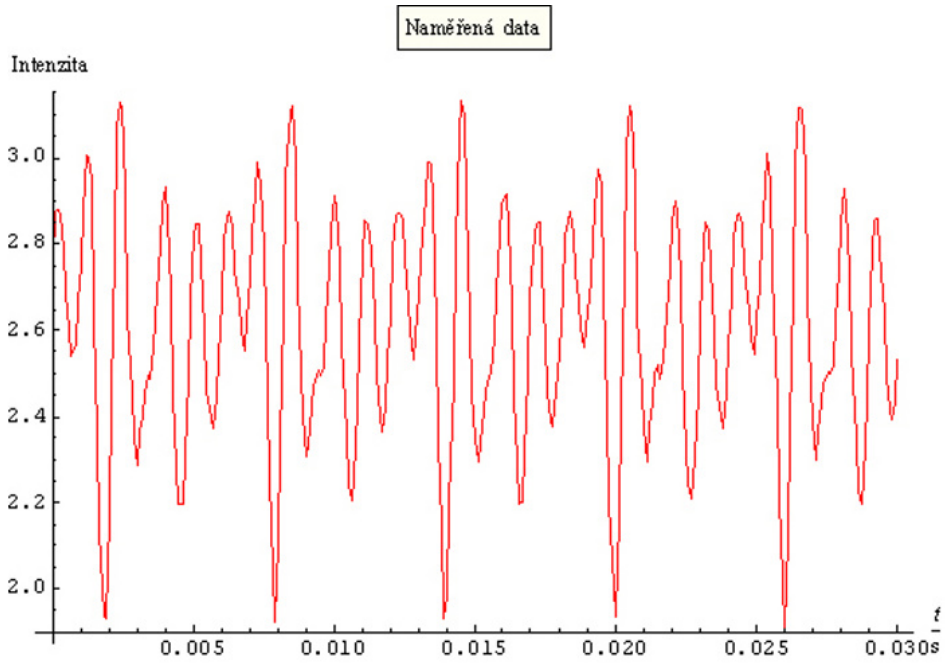
[7] notebook programu Mathematica, ve kterém ze data dále zpracovat

Přílohy

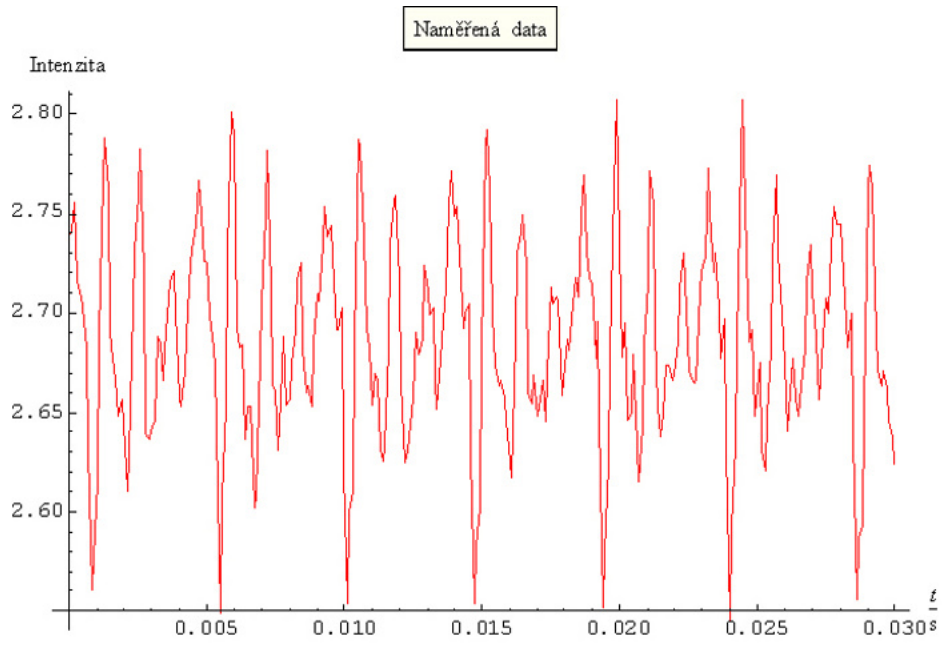
Na dalších obrázcích jsou uvedeny grafy závislostí intenzity na čase a frekvenční analýza pro samohlásky *a*, *e*, *i*, *o* a *u* (v tomto pořadí). Vždy jsou za sebou uvedeny charakteristiky moje a Šárky Lorencové, žákyně SPŠST Panská v Praze.

Z dále uvedených frekvencí je patrné, že pro většinu samohlásek má základní frekvence mého hlasu hodnotu 166 Hz a základní frekvence hlasu Šárky je 199 Hz. Podobné frekvence hlasu u mužů a žen jsou v souladu s výzkumy fyziologické akustiky.

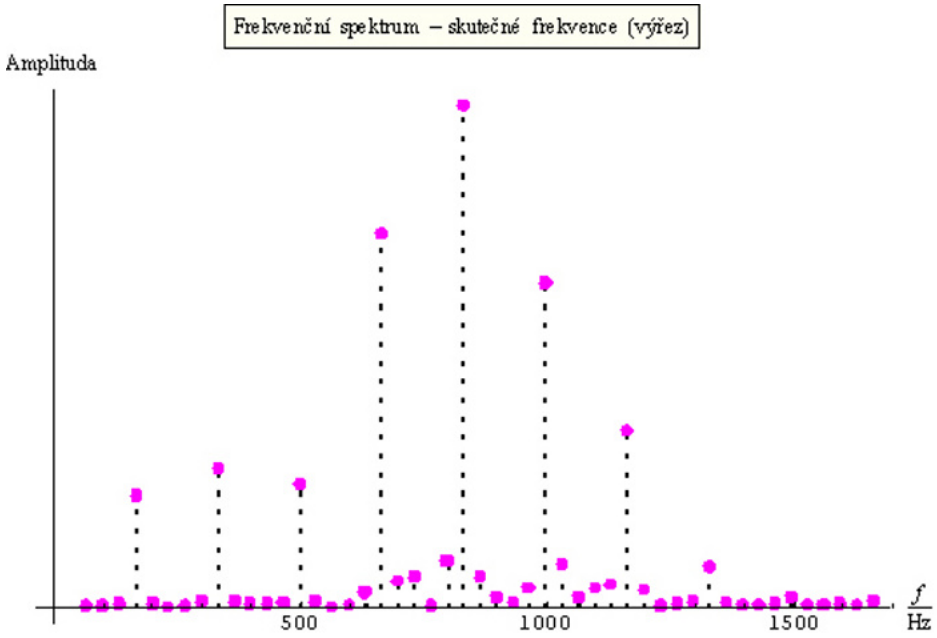
Samohláska *a*



obr. 2

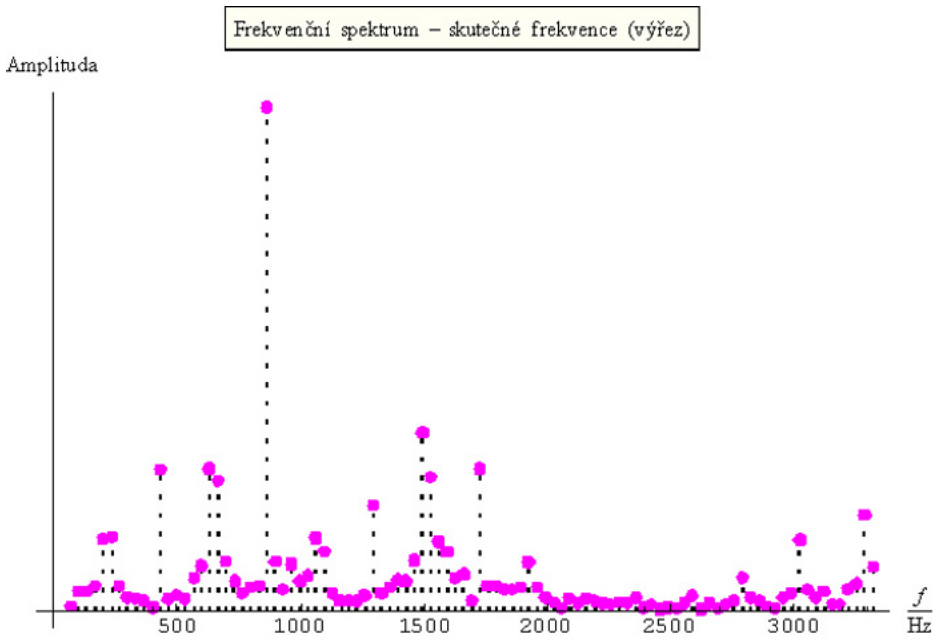


obr. 3



obr. 4

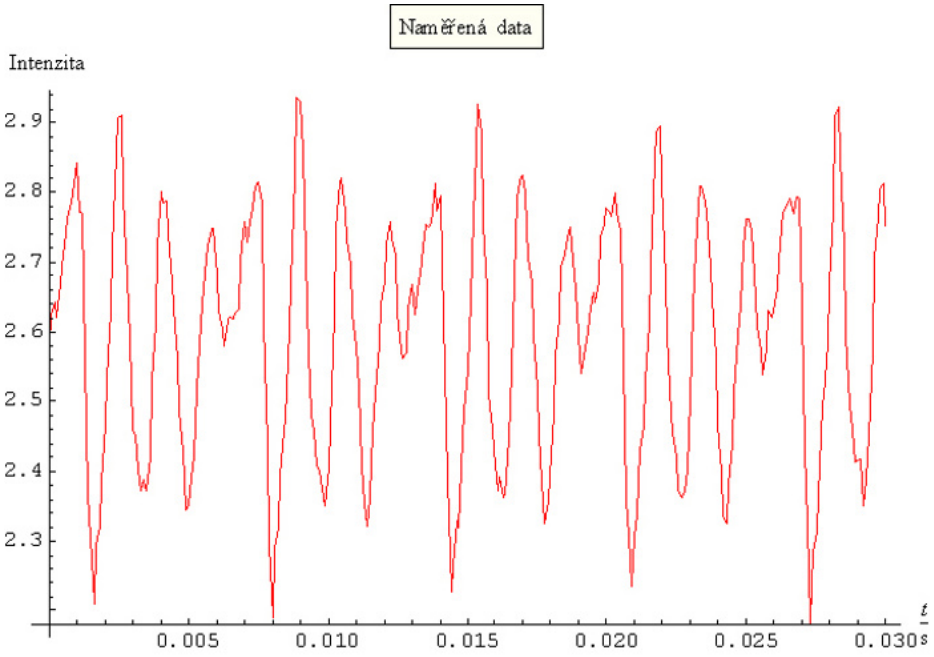
Zastoupené frekvence:
166,1 Hz; 332,2 Hz; 498,3 Hz; 664,5 Hz; 830,6 Hz; 996,7 Hz; 1162,8 Hz



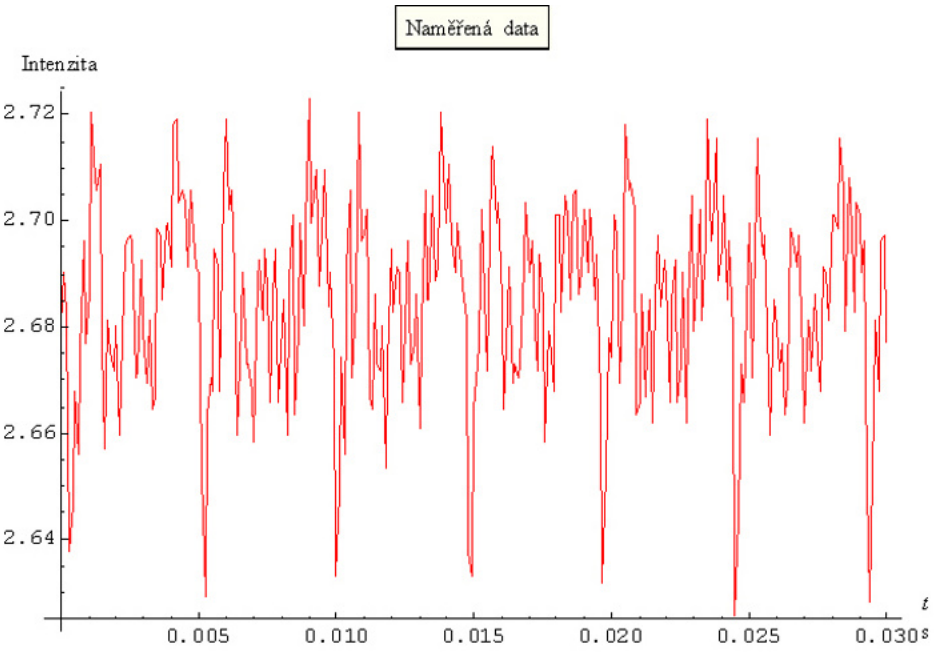
obr. 5

Zastoupené frekvence:
199,3 Hz; 232,6 Hz; 431,9 Hz; 631,2 Hz; 664,5 Hz; 863,8 Hz; 1063,1 Hz; 1295,7 Hz; 1495, Hz; 1528,2 Hz; 1561,5 Hz; 1727,6 Hz; 3023,3 Hz; 3289 Hz

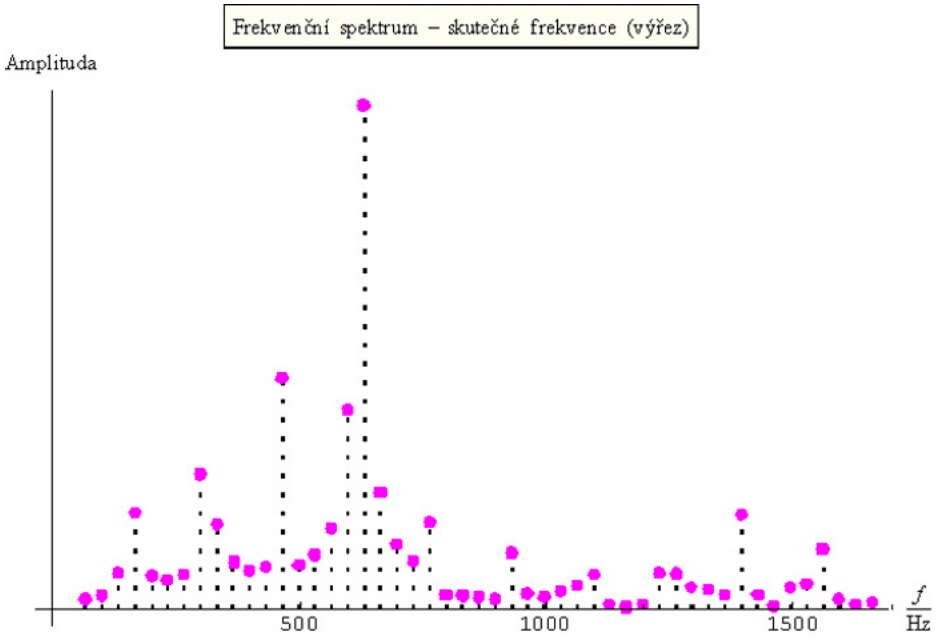
Samohláska *e*



obr. 6

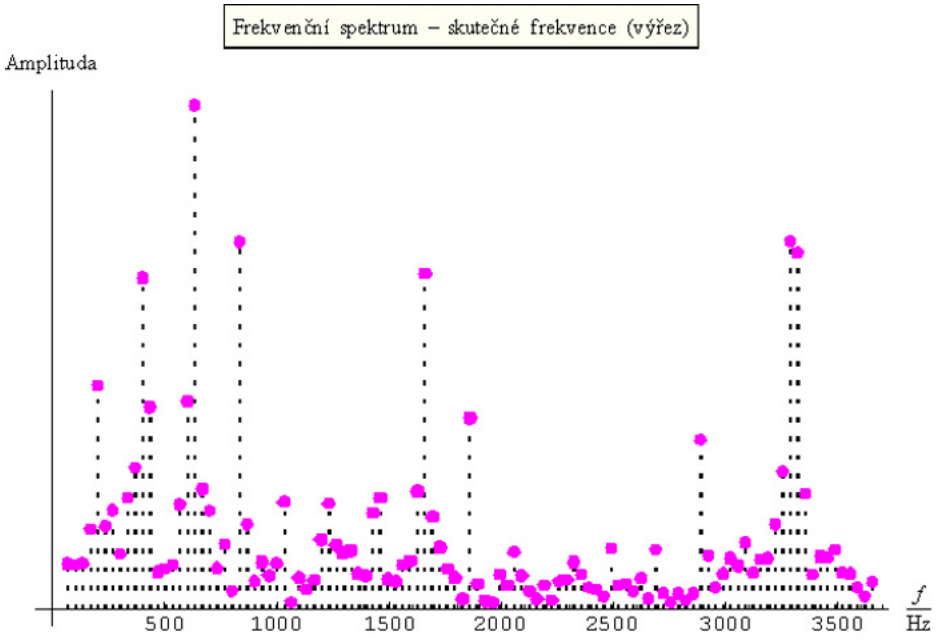


obr. 7



obr. 8

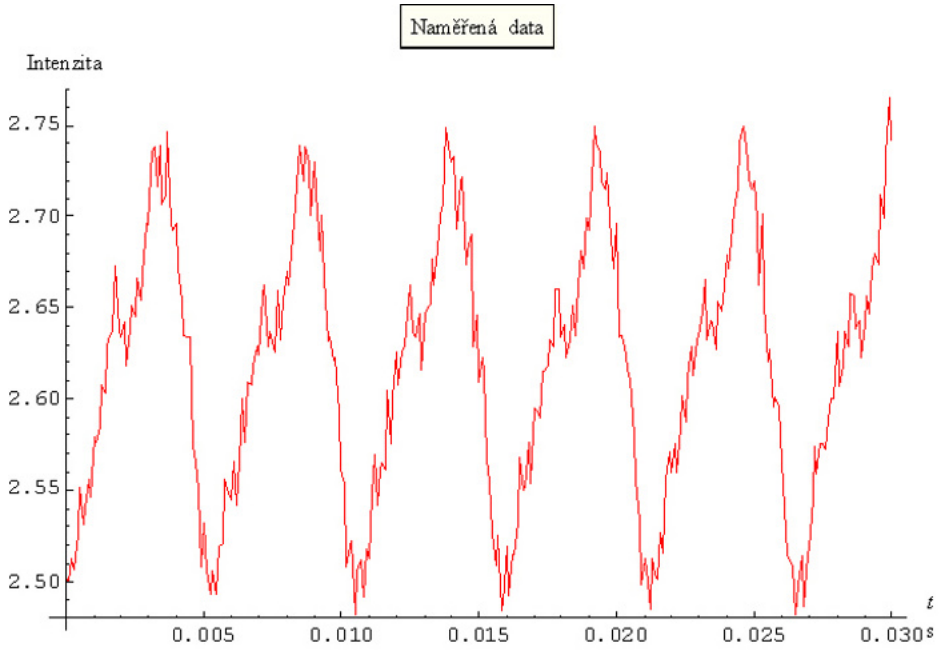
Zastoupené frekvence:
166,1 Hz; 299, Hz; 465,1 Hz; 598, Hz; 631,2 Hz; 664,5 Hz; 1395,3 Hz



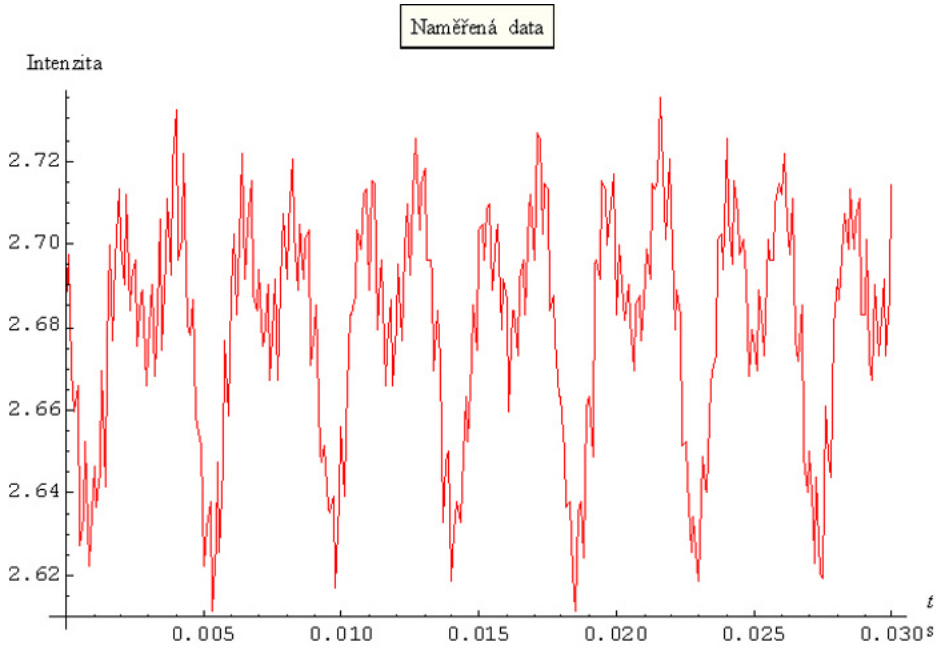
obr. 9

Zastoupené frekvence:
199,3 Hz; 398,7 Hz; 431,9 Hz; 598, Hz; 631,2 Hz; 830,6 Hz; 1661,1 Hz; 1860,5 Hz; 2890,4 Hz; 3289, Hz; 3322,3 Hz

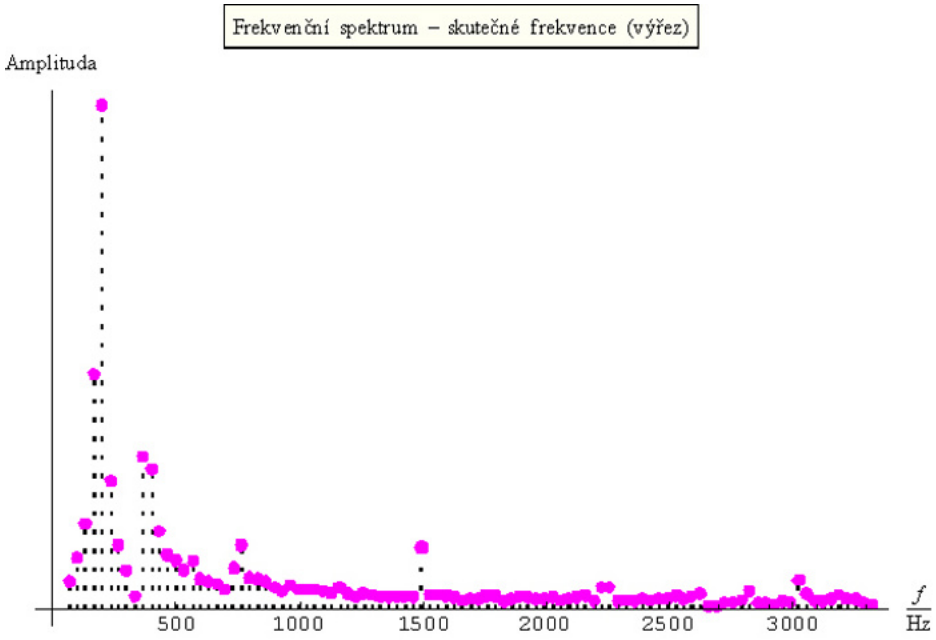
Samohláska *i*



obr. 10

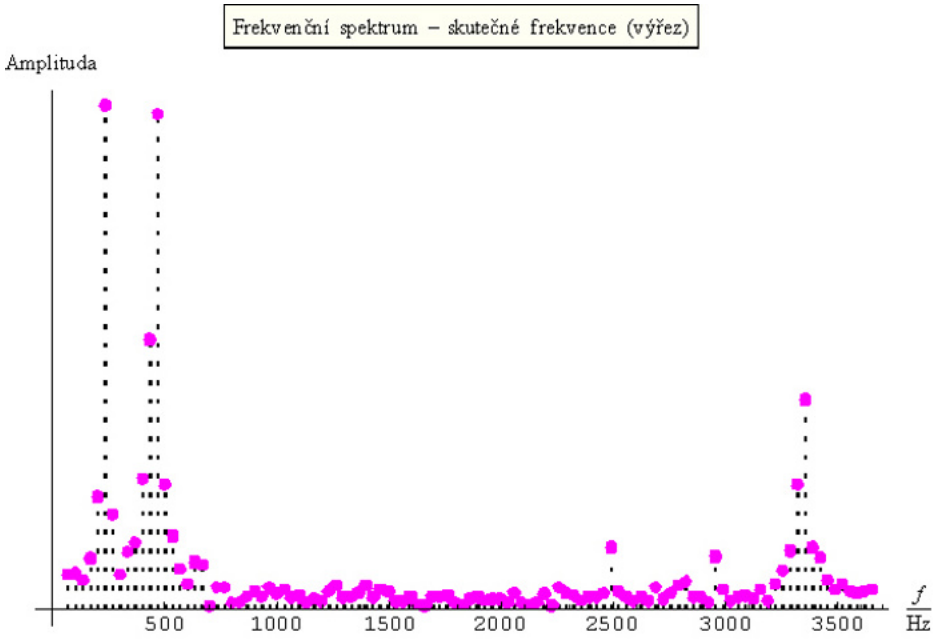


obr. 11



obr. 12

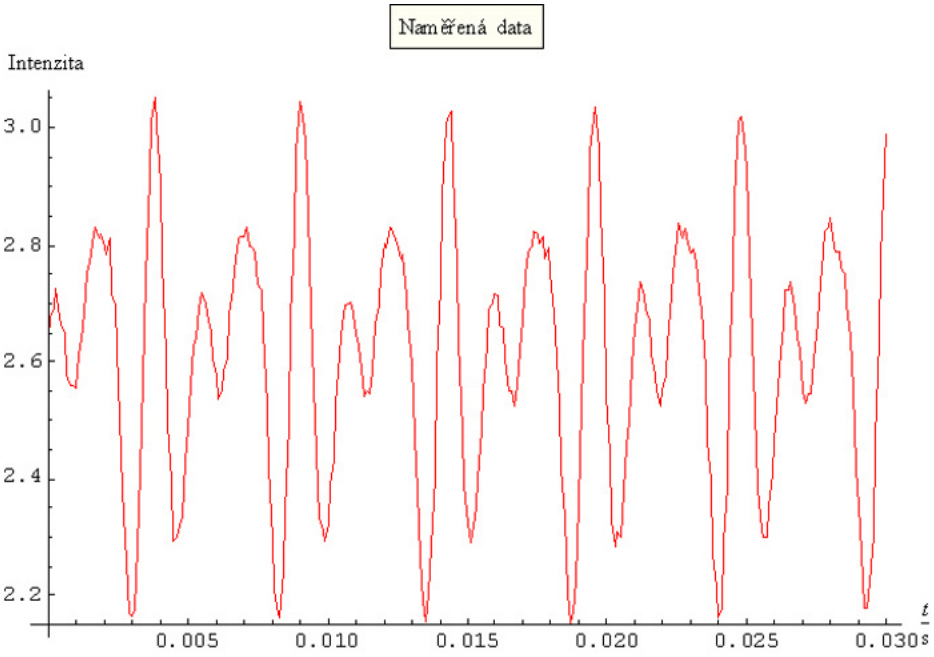
Zastoupené frekvence:
166,1 Hz; 199,3 Hz; 232,6 Hz; 365,4 Hz; 398,7 Hz



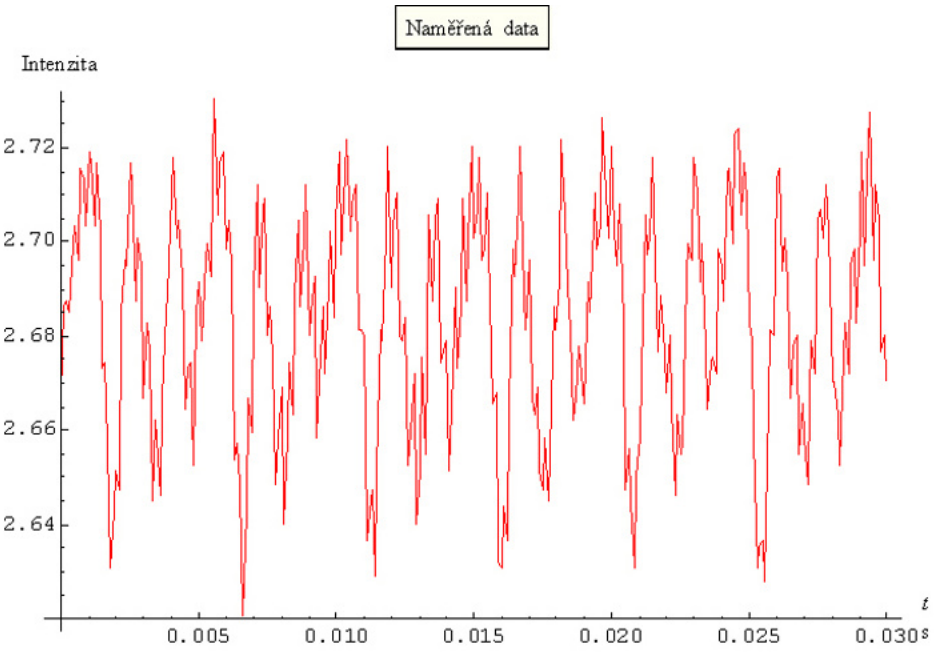
obr. 13

Zastoupené frekvence:
199,3 Hz; 232,6 Hz; 265,8 Hz; 398,7 Hz; 431,9 Hz; 465,1 Hz; 498,3 Hz; 3322,3 Hz; 3355,5 Hz

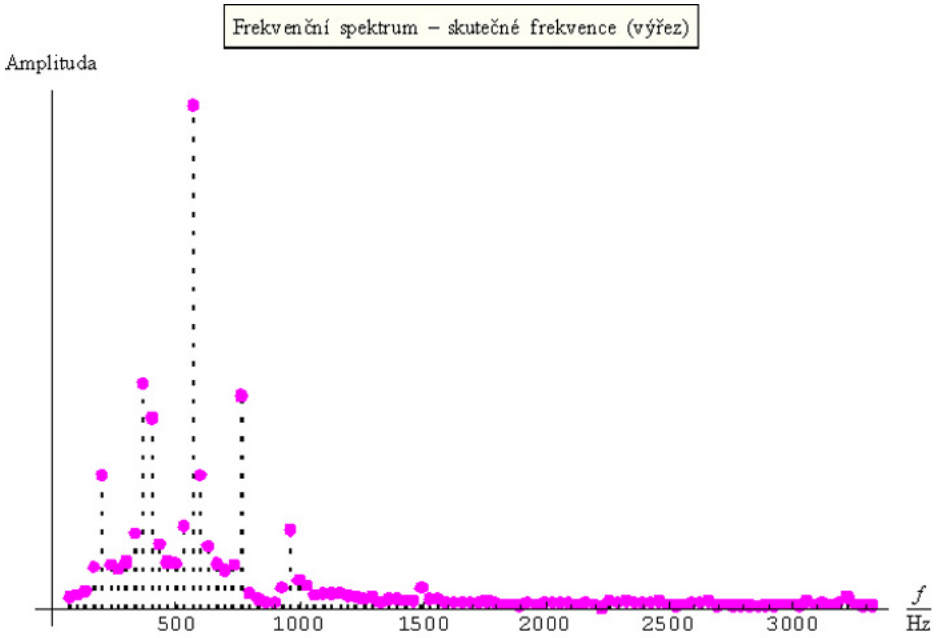
Samohláska *o*



obr. 14



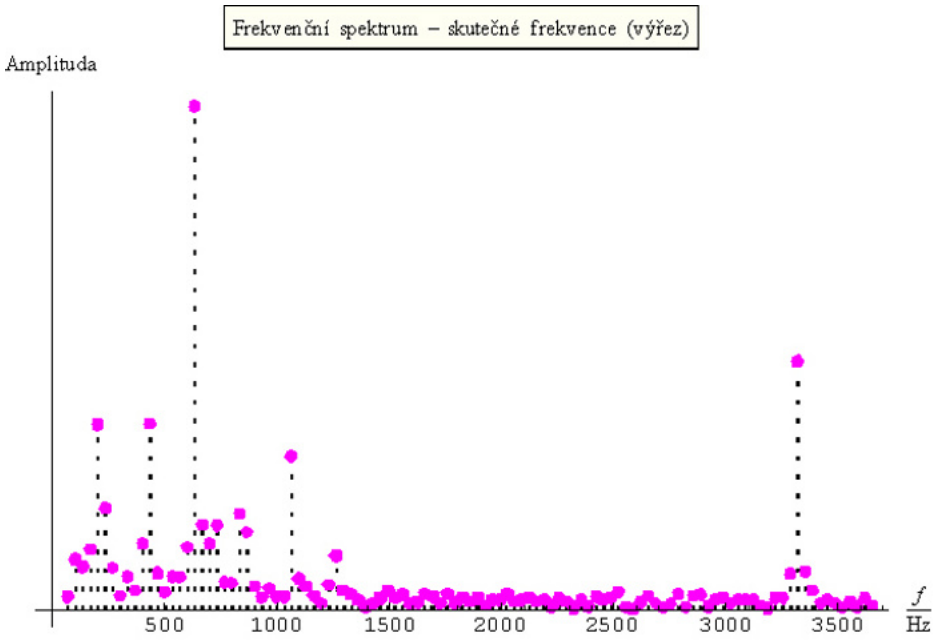
obr. 15



obr. 16

Zastoupené frekvence:

199,3 Hz; 332,2 Hz; 365,4 Hz; 398,7 Hz; 531,6 Hz; 564,8 Hz; 598, Hz; 764,1 Hz; 963,5 Hz

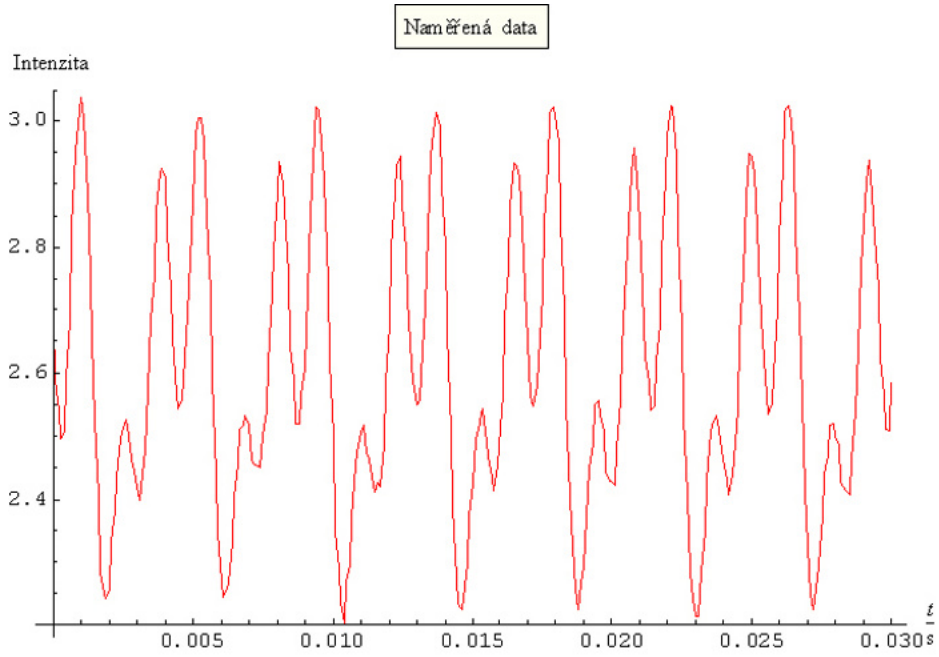


obr. 17

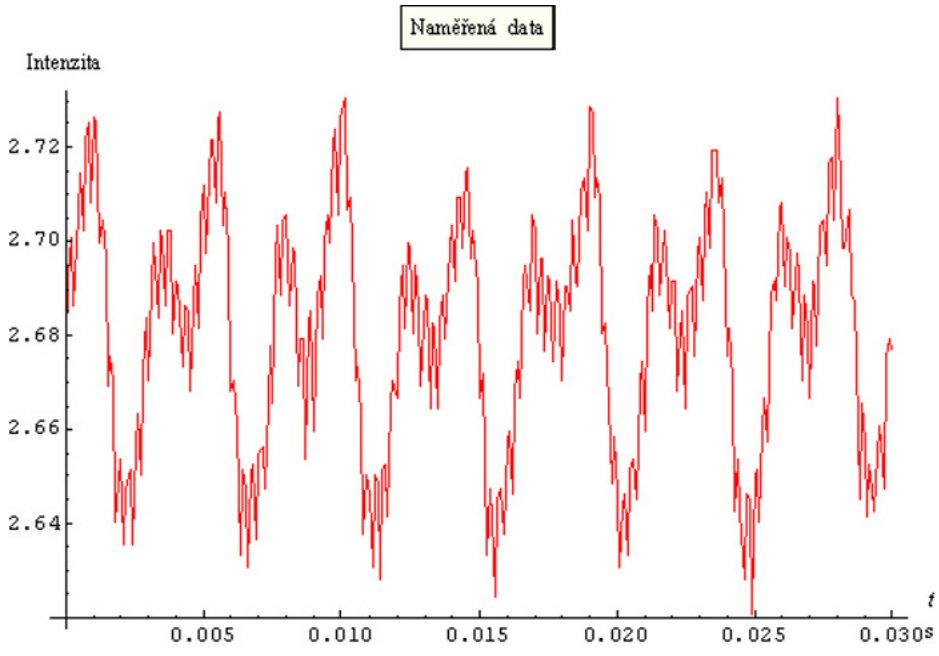
Zastoupené frekvence:

199,3 Hz; 431,9 Hz; 631,2 Hz; 1063,1 Hz; 3322,3 Hz

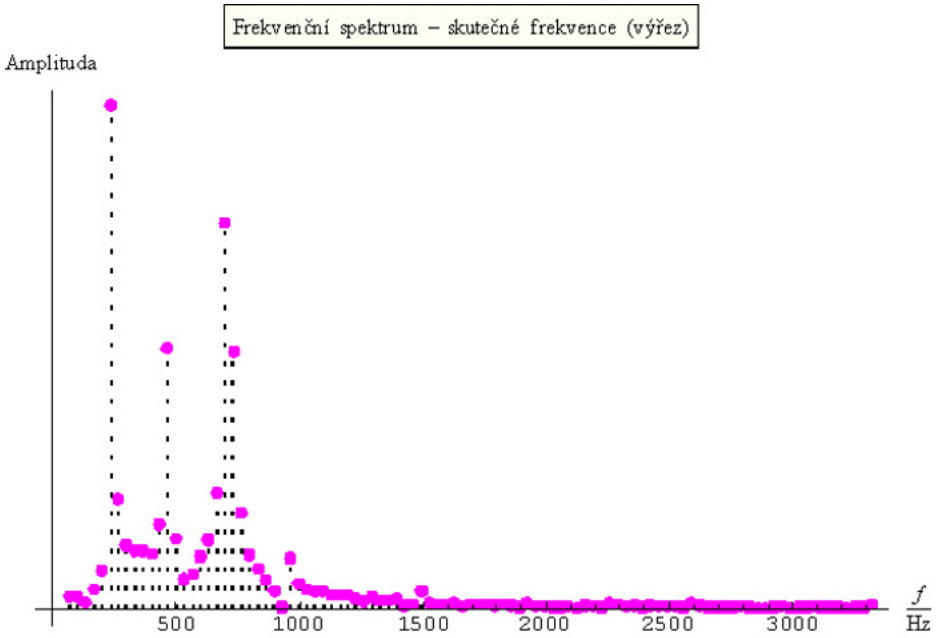
Samohláska *u*



obr. 18

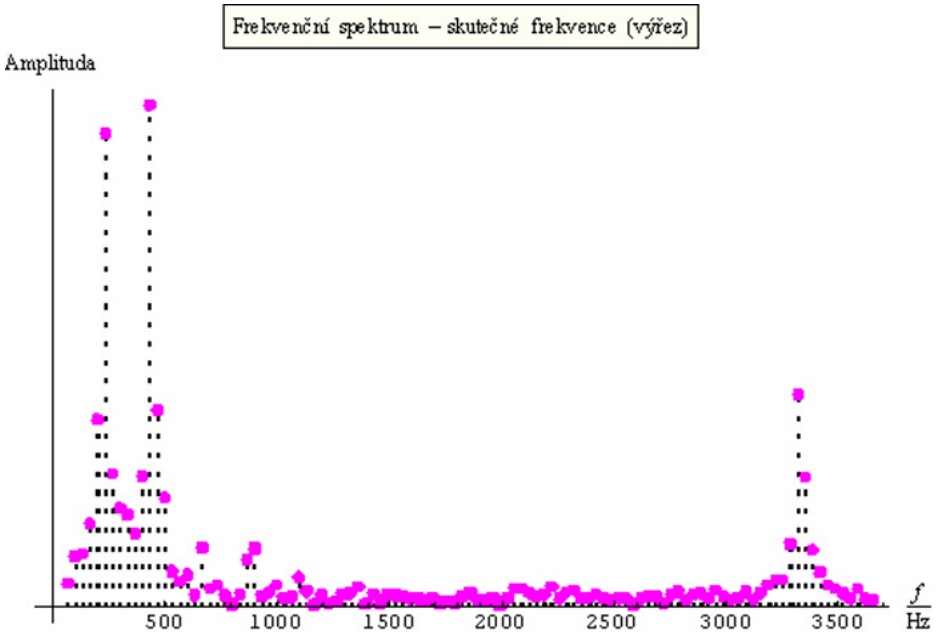


obr. 19



obr. 20

Zastoupené frekvence:
232,6 Hz; 265,8 Hz; 431,9 Hz; 465,1 Hz; 664,5 Hz; 697,7 Hz; 730,9 Hz; 764,1 Hz



obr. 21

Zastoupené frekvence:
166,1 Hz; 199,3 Hz; 232,6 Hz; 265,8 Hz; 299, Hz; 332,2 Hz; 398,7 Hz; 431,9 Hz; 465,1 Hz; 498,3 Hz; 3322,3 Hz; 3355,5 Hz