



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ZÁKLADNÍ AKUSTICKÉ VLASTNOSTI HLASU A ŘEČI



JAN G. ŠVEC ¹

¹ Katedra experimentální fyziky, Př.F. UP OLOMOUC



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ČTYŘI ZÁKLADNÍ VLASTNOSTI ZVUKU:

1) Periodicita: Základní frekvence (f_0)

2) Amplituda, Energie: Hladina akustického tlaku (SPL)

3) Barva: Spektrální rozložení

4) Trvání v čase



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ZÁKLADNÍ VLASTNOSTI ZVUKU

OBJEKTIVNÍ → **SUBJEKTIVNÍ (vnímané)**

Základní frekvence (f_0)



Výška

Intenzita, hladina akust. tlaku



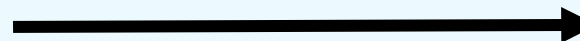
Hlasitost

Spektrum



Barva

Trvání v čase



Trvání v čase



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ZÁKLADNÍ FREKVENCE HLASU (f_0):

Měřena v hertzích [Hz]

Inverzně vztažena k periodě kmitání hlasivek T_0 , měřena

v sekundách [s]:

$$f_0 = 1/T_0$$

Subjektivně je vnímána jako výška hlasu (angl. pitch).

*Výška hlasu je většinou identifikována pomocí hudebních
tónů.*



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



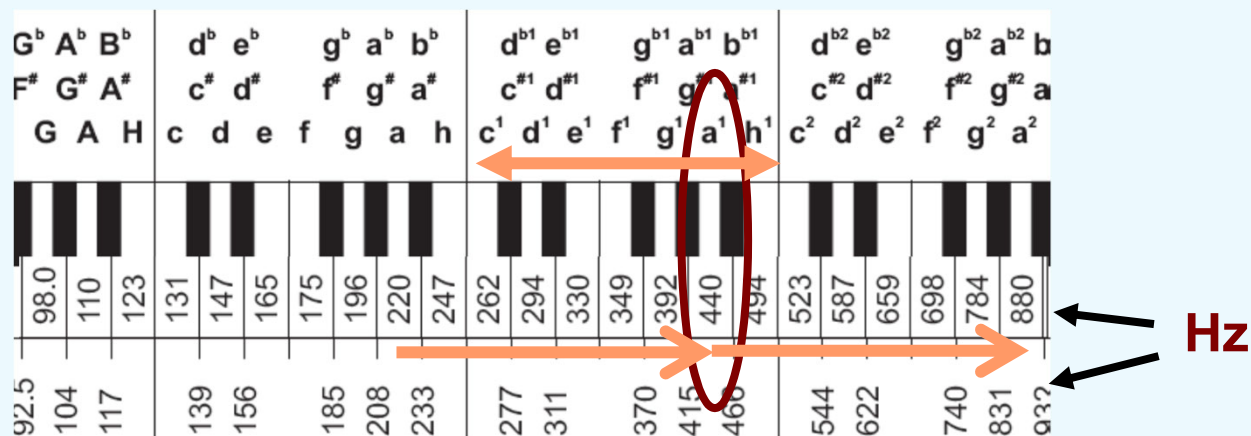
OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

FREKVENCE VERSUS HUDEBNÍ TÓNY:

Základní frekvence mohou být vztaženy k tónům na klavírní klaviatuře.

Je 12 půltónů na oktávu.



Referenční tón je komorní “a¹” s frekvencí 440 Hz

V mezinárodním systému je tento tón značen jako A₄.

Zvýšení o oktávu znamená zdvojnásobení frekvence v Hz.

Mezinárodní tónové značení začíná na tónu C₀ [16.352 Hz], který je 4 oktávy a 9 půltónů (t.j., 57 půltónů) níže než A₄.

MIDI norma začíná ještě o oktávu níže, na tónu C₋₁ [8.1758 Hz], tj. 69 půltónů pod A₄.

Midi tóny jsou logaritmicky vztaženy k frekvenci v hertzích: $MIDI = 69 + 12 \log_2(f_0/440)$. [Fechner-Weberův zákon]



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

FREKVENCE VERSUS TÓNY: Matematické vyjádření:

Zvýšení o oktávu znamená násobení frekvence číslem 2:

$$f_2 = 2 * f_1 \text{ [Hz]}$$

Jaký je faktor násobení x abychom zvýšili frekvenci f_1 o půltón?

Uvažujme, že oktáva má 12 půltónů, tedy platí:

$$f_2 = 2 * f_1 = x * x * x * x * x * x * x * x * x * x * x * x * f_1 = x^{12} * f_1 \text{ [Hz]}$$

Proto : $x^{12} = 2$ a tedy $x = 2^{1/12}$

Pro zvýšení frekvence f_1 o n půltónů platí:

$$f_2 = 2^{n/12} * f_1 \text{ [Hz]}$$

Pro snížení frekvence f_1 o n půltónů platí:

$$f_2 = 2^{(-n/12)} * f_1 \text{ [Hz]}$$



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

FREKVENCE VERSUS TÓNY: Matematické vyjádření:

Referenční frekvence 440 Hz je přiřazena MIDI normou pŕtónu č. 69.

Pro obdržení frekvence f_N N-tého pŕtónu podle MIDI normy platí vztah:

$$f_N = 2^{[(N-69) / 12]} * 440 \text{ [Hz]}$$

Pro určení MIDI pŕtónu N ze známé frekvence v hertzích potřebujeme odvodit inverzní

logaritmický vztah. Platí:

$$(N-69)/12 = \log_2(f_N/440), \text{ tedy:}$$

$$N = 12 * \log_2(f_N/440) + 69$$

Pro náhradu logaritmu se základem 2 (\log_2) běžným dekadickým logaritmem (\log_{10}),

platí:

$$\log_2(y) = \log_{10}(y) / \log_{10}(2)$$

Výsledný vztah tedy je:

$$N = 12 * \log_2(f_N/440) + 69 = [12 * \log_{10}(f_N/440)] / [\log_{10}(2)] + 69 \text{ [MIDI]}$$

FREKVENCE VERSUS VÝŠKA TÓNU:

ČESKÉ ZNAČENÍ

| Oktávy | sub-kontra | kontra | velká | malá | jednočárková | dvoučárková | tříčárková | čtyřčárková | | | | | | | | |
|----------------|----------------------------------|---|---|---|---|---|---|--|--|--|---|------------------------------------|---|------------------------------------|---|-----|
| Tóny | B ₂ A ₂ | D ^b , E ^b , C [#] , D ₁ | G ^b , A ^b , B ₁ , C [#] , D [#] | D ^b , E ^b , C [#] , D [#] | G ^b , A ^b , B ^b , F [#] , G [#] , A [#] | c [#] , d [#] , e [#] , f [#] , g [#] , a [#] | c ^{#1} , d ^{#1} , e ^{#1} , f ^{#1} , g ^{#1} , a ^{#1} | c ^{#2} , d ^{#2} , e ^{#2} , f ^{#2} , g ^{#2} , a ^{#2} | d ^{#3} , e ^{#3} , f ^{#3} , g ^{#3} , a ^{#3} | d ^{#4} , e ^{#4} , f ^{#4} , g ^{#4} , a ^{#4} | g ^{b4} , a ^{b4} , b ^{b4} | | | | | |
| Tóny | A ₂ H ₂ | C ₁ , D ₁ , E ₁ , F ₁ , G ₁ , A ₁ , H ₁ | C, D, E, F, G, A, H | c, d, e, f, g, a, h | c ¹ , d ¹ , e ¹ , f ¹ , g ¹ , a ¹ , h ¹ | c ² , d ² , e ² , f ² , g ² , a ² , h ² | c ³ , d ³ , e ³ , f ³ , g ³ , a ³ , h ³ | c ⁴ , d ⁴ , e ⁴ , f ⁴ , g ⁴ , a ⁴ , h ⁴ , c ⁵ | | | | | | | | |
| Frekvence [Hz] | | 32.7, 34.6, 36.7, 38.9, 41.2 | 43.7, 46.2, 49.0, 51.9, 55.0, 58.3, 61.7 | 65.4, 69.3, 73.4, 77.8, 82.4 | 87.3, 92.5, 98.0, 104, 110, 117, 123 | 131, 139, 147, 156, 165 | 175, 185, 196, 208, 220, 233, 247 | 262, 277, 294, 311, 330 | 349, 370, 392, 415, 440, 466, 494 | 523, 544, 587, 622, 659 | 698, 740, 784, 831, 880, 932, 988 | 1046, 1109, 1175, 1245, 1319 | 1397, 1480, 1568, 1661, 1760, 1865, 1976 | 2093, 2217, 2349, 2489, 2637 | 2794, 2960, 3136, 3322, 3520, 3729, 3951, 4186 | |
| MIDI frekvence | 21, 23 | 24, 26, 28 | 29, 31, 33, 35 | 36, 38, 40 | 41, 43, 45, 47 | 48, 50, 52 | 53, 55, 57, 59 | 60, 62, 64 | 65, 67, 69, 71 | 72, 74, 76 | 77, 79, 81, 83 | 84, 86, 88 | 89, 91, 93, 95 | 96, 98, 100 | 101, 103, 105, 107 | 108 |
| Tóny | A ₀ , B ₀ | C ₁ , D ₁ , E ₁ , F ₁ , G ₁ , A ₁ , B ₁ | C ₂ , D ₂ , E ₂ , F ₂ , G ₂ , A ₂ , B ₂ | C ₃ , D ₃ , E ₃ , F ₃ , G ₃ , A ₃ , B ₃ | C ₄ , D ₄ , E ₄ , F ₄ , G ₄ , A ₄ , B ₄ | C ₅ , D ₅ , E ₅ , F ₅ , G ₅ , A ₅ , B ₅ | C ₆ , D ₆ , E ₆ , F ₆ , G ₆ , A ₆ , B ₆ | C ₇ , D ₇ , E ₇ , F ₇ , G ₇ , A ₇ , B ₇ | C ₈ | | | | | | | |
| Oktávy | nultá | první | druhá | třetí | čtvrtá | pátá | šestá | sedmá | | | | | | | | |

MEZINÁRODNÍ ZNAČENÍ

Poznámky:

Standardní, rovnoměrně temperované ladění má jako referenční tón 440 Hz (komolí a) a uvažuje 12 stejných půltónů na oktávu.

Mezinárodní značení tónů začíná na tónu C₀, o frekvenci 16.352Hz (57 půltónů pod komorním a)

MIDI standard začíná na tónu C₁, o frekvenci 8.1758 Hz (69 půltónů pod komorním a). MIDI frekvence: $d = 69 + 12 \cdot \log_2(f/440)$



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ZÁKLADNÍ FREKVENCE HLASU:

Je určena kmitáním hlasivek

PRAVIDELNÉ KMITY: Základní frekvenci lze určit z počtu kmitů
hlasivek uskutečněných za 1 sekundu [měřeno v Hz]

NEPRAVIDELNÉ KMITY: Základní frekvence není dobře
definovatelná, někdy je ji nemožné změřit (chraptivé hlasy, žádný
specifický tón)



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ZÁKLADNÍ FREKVENCE HLASU:

Změny f_0 lze zviditelnit pomocí tzv. úzkopásmového spektrogramu.

Co je to spektrogram?



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ZVUKOVÁ SPEKTROGRAFIE (SONAGRAFIE)

Vyvinuta zhruba kolem roku 1940 v
Bellových laboratořích v USA
Byla nazývána „viditelná řeč“ („visible
speech“).

Zobrazuje frekvenční spektrum zvuku
(hlasu / řeči) v čase. (V podstatě se
jedná o 3D spektrální analýzu)

Signál je analyzován pomocí filtrů s
nastavitelnou šířkou propustného
pásma.

Je hodnocena intenzita signálu v
různých frekvenčních pásmech a ta
se zobrazí mírou zčernání.

Nyní je běžně dostupná na bázi FFT v
různých počítačových softwarech pro
akustická měření.

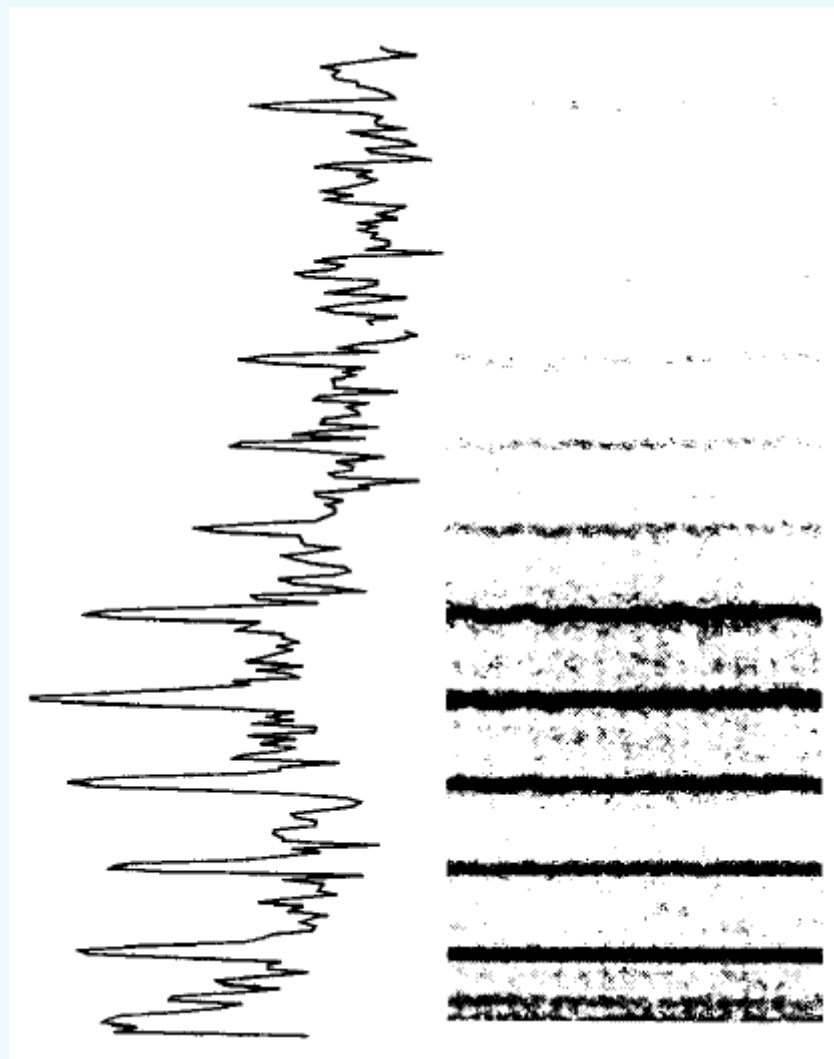


Fig. 7.27: Baken RJ, and Orlikoff RF: Clinical measurement of speech and voice, San Diego, CA: Singular Publishing Group, 2000.



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ZVUKOVÁ SPEKTROGRAFIE (SONAGRAFIE)

Vyvinuta zhruba kolem roku 1940 v Bellových laboratořích v USA

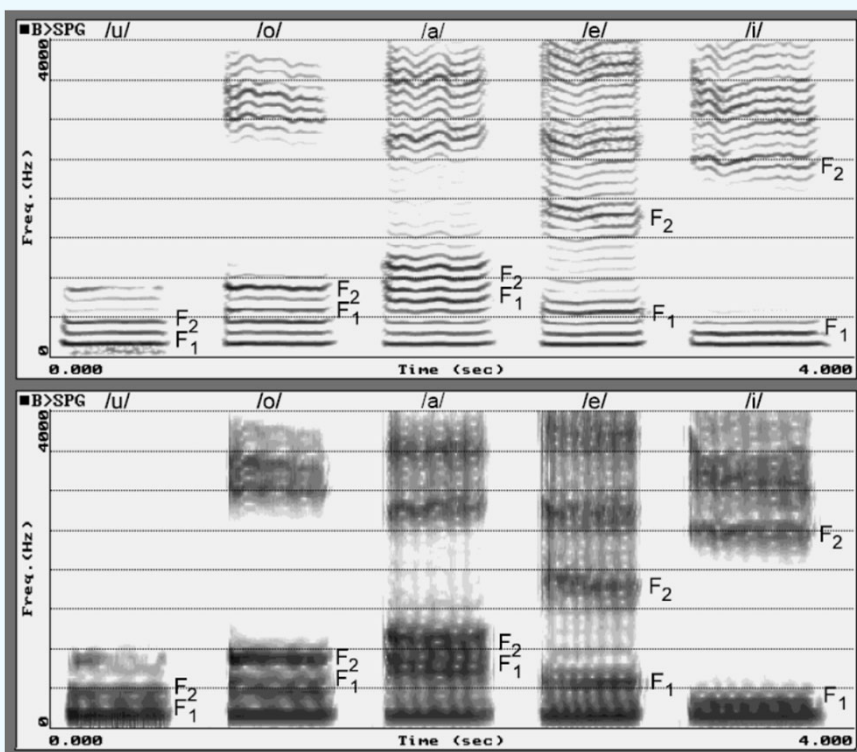
Byla nazývána „viditelná řeč“ („visible speech“).

Zobrazuje frekvenční spektrum zvuku (hlasu / řeči) v čase. (V podstatě se jedná o 3D spektrální analýzu)

Signál je analyzován pomocí filtrů s nastavitelnou šířkou propustného pásma.

Je hodnocena intenzita signálu v různých frekvenčních pásmech a ta se zobrazí mírou zčernání.

Nyní je běžně dostupná na bázi FFT v různých počítačových softwarech pro akustická měření.



2 typy spektrogramu:

a) úzkopásmový

- šířka propustného pásma Δf filtrů je řádově desítky hertzů,
- vhodný pro sledování změn základní frekvence signálu
- vysoké frekvenční rozlišení Δf , pomalé časové rozlišení Δt

b) širokopásmový

- šířka propustného pásma Δf filtrů je řádově stovky hertzů,
- vhodné pro zobrazování formantů
- Rychlé časové rozlišení ale malé frekvenční rozlišení
- Objevují se zde vertikální čáry, které jsou synchronizované s jednotlivými periodami kmitů

Platí vztah mezi časovým a frekvenčním rozlišením: $\Delta t = 1 / \Delta f$.

(Čím lepší/rychlejší časové rozlišení, tím menší/horší frekvenční rozlišení).

Švec, Obr. 2.1.
Spektrogramy
českých
samohlásek /u/,
/o/, /a/, /e/, /i/.
Nahoře
úzkopásmový
spektrogram, dole
širokopásmový
spektrogram.



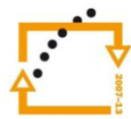
evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ZVUKOVÁ SPEKTROGRAFIE (SONAGRAFIE)

Terminologie:

Spektrografie = metoda používající spektra a spektrogramy

Spektrogram = graf, který zobrazuje spektrum signálu v čase

Zvukový spektrogram = sonagram = graf, který zobrazuje spektrum zvuku v čase

Spektrograf = přístroj, který vytváří spektrogram



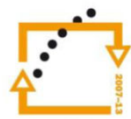
evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

DEMONSTRACE zvuková spektrografie

VOCE VISTA,

PRAAT



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ZVUKOVÁ SPEKTROGRAFIE:

CVIČENÍ

Proved' spektrografickou analýzu pomocí software Praat.

Urči nejnižší a nejvyšší základní frekvenci své řeči pomocí úzkopásmového spektrogramu.

Urči formanty své samohlásky [a:] pomocí širokopásmového spektrogramu.



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



**OP Vzdělávání
pro konkurenceschopnost**

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ZÁKLADNÍ FREKVENCE HLASU:

Normativní data

Z-skóre



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

ZÁKLADNÍ FREKVENCE HLASU:

U hlasu a řeči je přesné měření základní frekvence netriviální záležitostí.

Kmitání hlasivek není zcela pravidelné.

I u zdravých čistých hlasů vykazuje perioda kmitání hlasivek variabilitu –

hlas je proto označován jako **kvaziperiodický.**

Čím více je hlas porušen, tím více jsou kmity hlasivek nepravidelné.

Odchytky od ideálně periodických, předvídatelných kmitů se nazývají

PERTURBACE.



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

PERTURBACE HLASU:

ZÁKLADNÍ VELIČINY PERTURBACÍ:

Jitter – perturbace periody kmitů (u normálního hlasů má být pod 1%)

Shimmer – perturbace amplitudy kmitů
(u akustického signálu normálního hlasu má být pod 4%)



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

PŘEHLED PERTURBACÍ HLASU A JEJICH MĚŘENÍ:

Workshop on Acoustic Voice Analysis

SUMMARY STATEMENT
BY INGO R. TITZE, PH.D.

1995

NCVS

National Center for Voice and Speech

Tato publikace zavádí 3 typy signálů

http://www.ncvs.org/products_freebooks.html



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

TŘI TYPY SIGNÁLŮ / KMITŮ:

Kymografické záznamy kmitů hlasivek

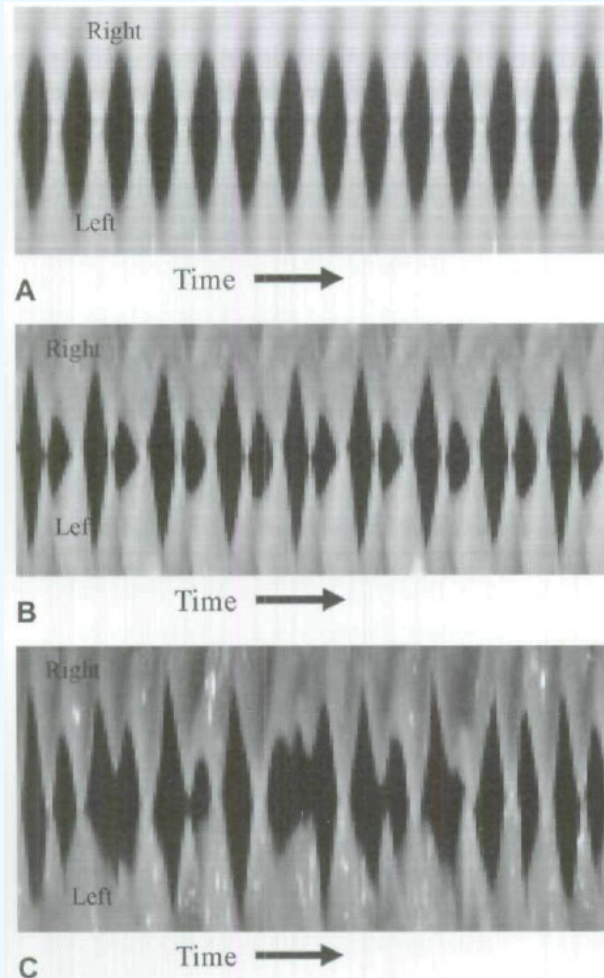


Fig 2. Digital kymography images. A) Sequence of type 1 vocal fold vibrations. B) Sequence of type 2 vocal fold vibrations. C) Sequence of type 3 vocal fold vibrations.

Zhang Y et al. Typing vocal fold vibratory patterns in excised larynx experiments via digital kymography. *Ann.Otol.Rhinol.Laryngol.* 118 (8):598-605, 2009.

TYP 1: téměř pravidelné kmity které nevykazují žádné kvalitativní změny v analyzovaném segmentu hlasu; pokud jsou přítomny modulační nebo subharmonické frekvence, jejich energie jsou o řád nižší než energie základní frekvence.

TYP 2: kmity vykazující kvalitativní změny (bifurkace) v analyzovaném segmentu hlasu, nebo kmity se subharmonickými či modulačními frekvencemi, jejichž energie se blíží energii základní frekvence; není zde jasně zřetelná jenom jedna jediná základní frekvence v rámci segmentu.

TYP 3: kmity bez zjevné periodické struktury.

I. R. Titze. *Workshop on acoustic voice analysis. Summary statement*, Denver, CO: National Center for Voice and Speech, 1995.



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

TŘI TYPY KMITŮ V AUDIO SIGNÁLU:

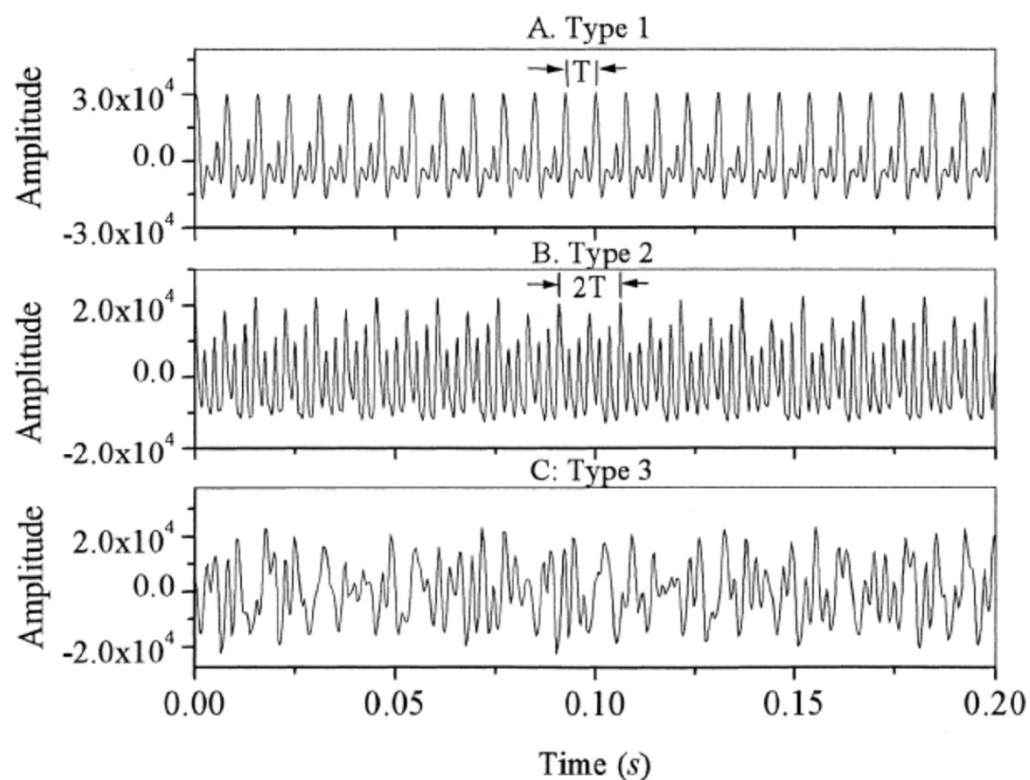


FIGURE 1. The typical waveforms of type 1, type 2, and type 3 voice signals.

J. J. Jiang, Y. Zhang, and C. McGilligan. Chaos in voice, from modeling to measurement. J.Voice 20 (1):2-17, 2006.

TYP 1:

Jasná periodicitu – f_0
Ize přibližně určit.

TYP 2:

Subharmonický signál,
komplexní periodicitu.
 f_0 analýza nedává
jednoznačné výsledky
– různé algoritmy
dávají různé výsledky.

TYP 3:

Nepravidelný signál. f_0
nelze spolehlivě určit.

**Smysluplnou f_0
analýzu lze provést
pouze na signálech
TYPU 1. (Titze 1995).**



evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

TŘI TYPY SIGNÁLŮ ve spektrogramu:

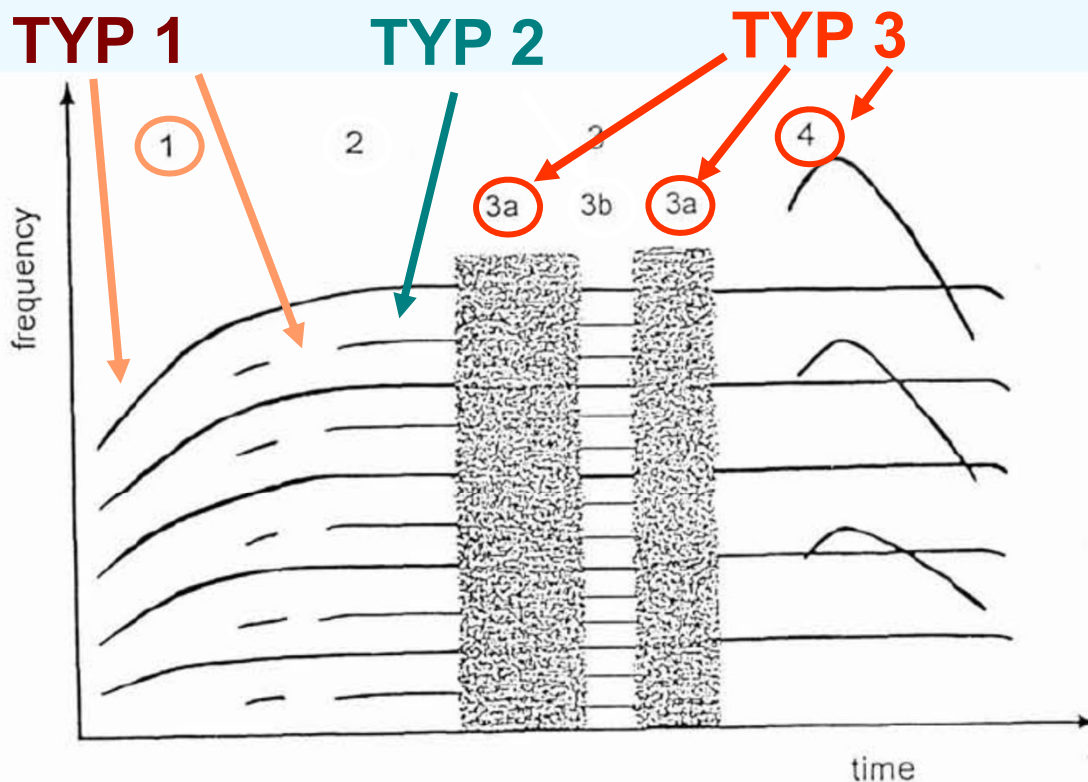


Fig. 2. Sketch of a narrow-band spectrogram displaying periodic phonation (1), subharmonic regimes (2), chaos (3a) interrupted by a periodic window (3b), and biphonation (4). Reproduced from Ref.[23] with permission.

Herzel H. Nonlinear dynamics of the voice: time series analysis, modeling and experiments. *Current Topics in Acoustical Research* 2:17-30, 1998.

1) (Kvazi/téměř)
periodická fonace
(TYP 1)

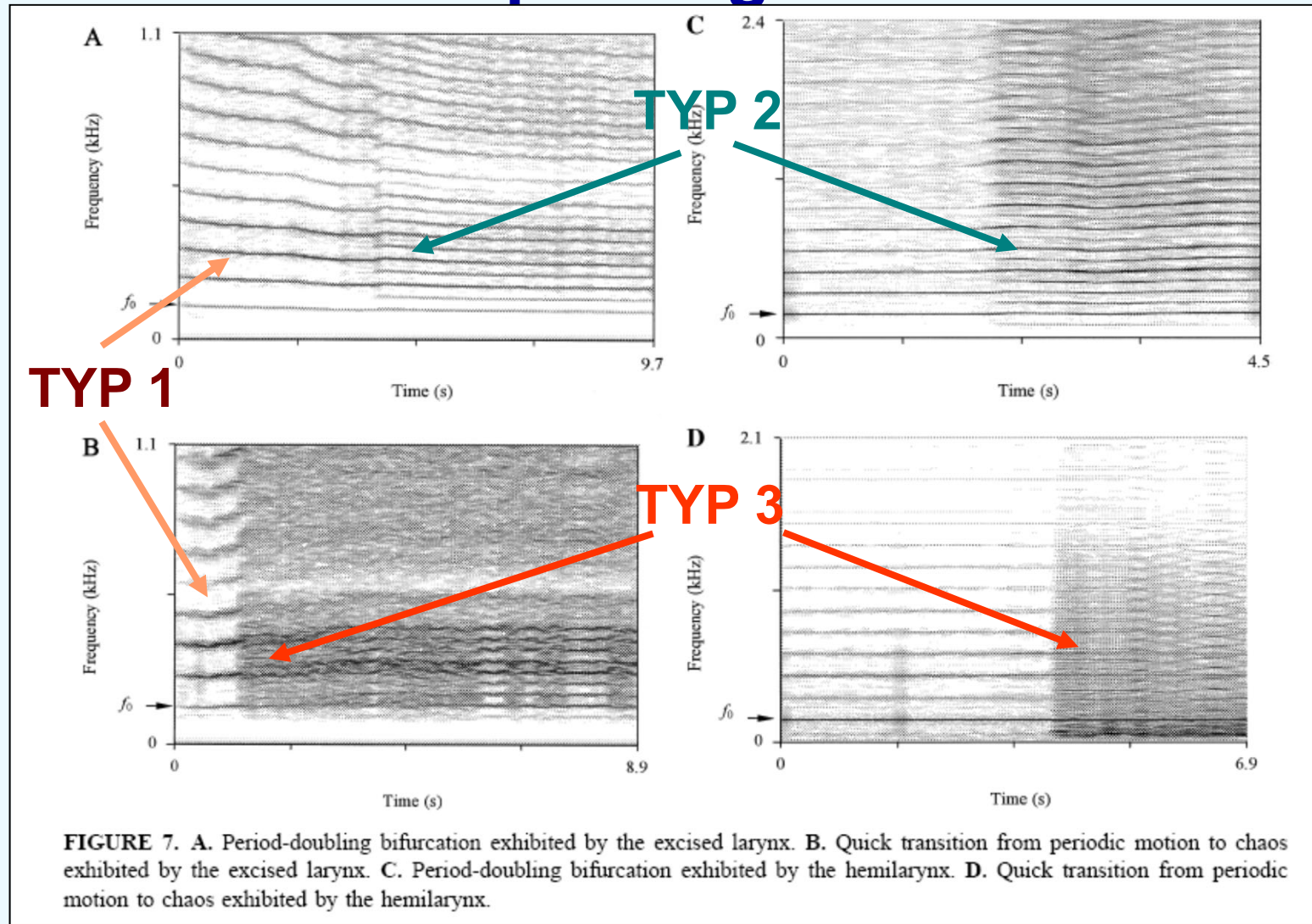
2) Subharmonická
fonace na frekvenci
 $f_{\text{SUB}} = f_0/2$
(TYP 2)

3a) Nepravidelná
fonace
(TYP 3)

3b) Subharmonická
fonace na frekvenci
 $f_{\text{SUB}} = f_0/3$
(TYP 2)

4) Bifonace
- dvě nezávislé
frekvence znějící
zároveň (TYP 3)

TŘI TYPY SIGNÁLŮ ve spektrogramu:





evropský
sociální
fond v ČR



EVROPSKÁ UNIE



MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



OP Vzdělávání
pro konkurenceschopnost

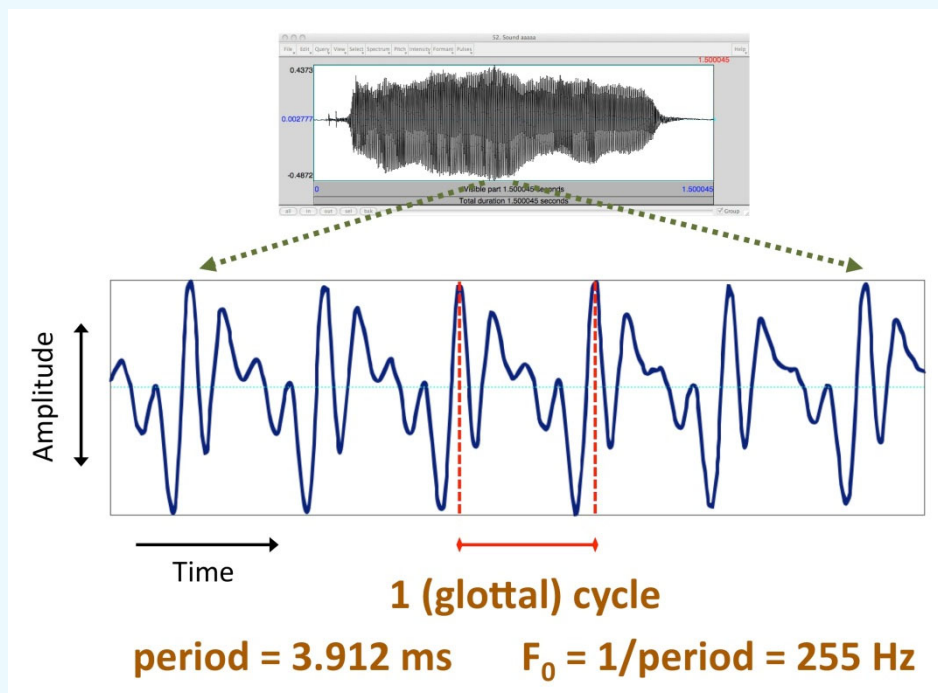
INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

Periodicita u hlasových signálů

- Perioda: trvání jednoho cyklu hlasivkových kmitů
- F_0 lze určit pouze u PERIODICKÝCH signálů
- Hlasové signály jsou kvaziperiodické
- Změny periody: **VIBRATO** (pravidelné), **JITTER** (nepravidelné)



Slide by C. Herbst





MINISTERSTVO ŠKOLSTVÍ,
MLÁDEŽE A TĚLOVÝCHOVY



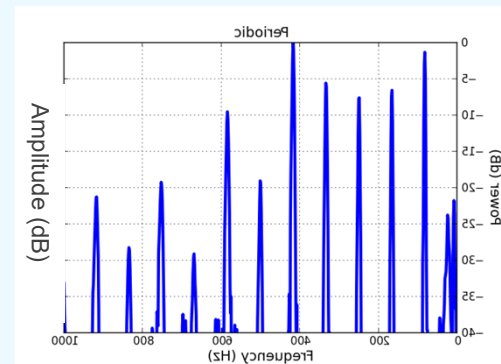
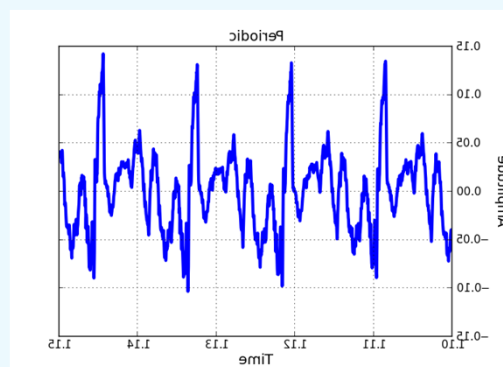
OP Vzdělávání
pro konkurenceschopnost

INVESTICE
DO ROZVOJE
VZDĚLÁVÁNÍ

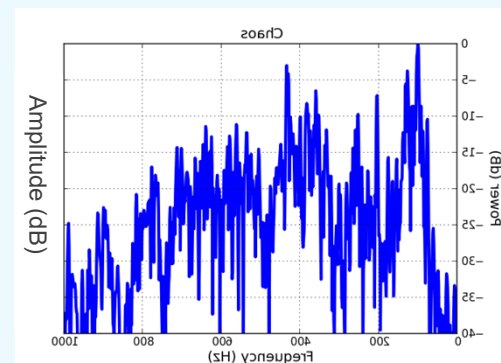
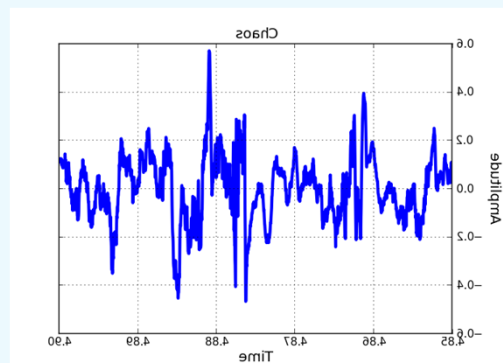
Pravidelné a nepravidelné kmity hlasivek



Pravidelné



Nepravidelné



Slide by C. Herbst



Spektrum hlasu: harmonická série

šumové složky,
hluk pozadí



Slide by C. Herbst

