

Multimediální systémy

07 - Zvuk

Osnova přednášky

- Zvuk – fyzikální podstata a vlastnosti
- Digitální zvuk
- Komprese, kodeky, formáty
- Zvuk v MMS

Přítomnost zvuku výrazně zvyšuje
efekt grafické prezentace v jakékoliv
podobě

ALE!

Zvuk

- Když v lese spadne strom a není v něm nikdo, kdo by to slyšel, vznikne zvuk?
- Zvuk = organizovaný pohyb molekul, vyvolaný určitým objektem, pohybující se v určitém médiu (vzduch, voda, ...)
- Zvuk = sluchový zážitek, který jsme schopni vnímat

Fyzikální podstata zvuku

- Pokud se objekt pohybuje sem a tam (vytváří vibrace), posune vzduch nacházející se přímo u něj
- Vytváří se tím vlna podobné soustředným kolům, které vytvoříme vhozením kamene do stojaté vody
- Částičky vzduchu pohybující se ve vlnách zapříčiňují kmitání ušního bubínku
- Kmitání je registrováno sérií kůstek (kladívko, kovadlinka, třmínek) a přenášeno do nervů vnitřního ucha
- To posílá impulsy do mozku, který je vnímá jako zvuk
- zvuk = „vlnění vzduchu“

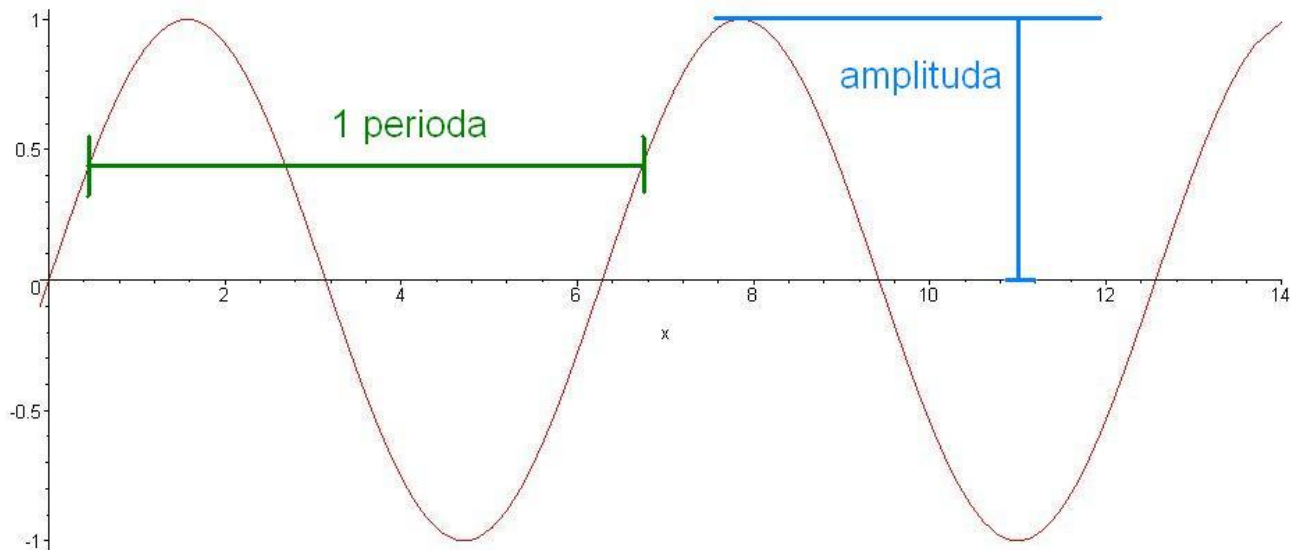


Fyzikální podstata zvuku

- *Proč vnímáme vlastní hlas jinak než jej vnímají všichni ostatní lidé?*

Fyzikální vlastnosti zvuku

- Perioda – doba „celé vlny“ – doba mezi dvěma vrcholy
- Frekvence – počet vrcholů vlny za 1 sekundu (Hz), člověk 20 – 20 000 Hz (s věkem se rozsah snižuje, i na 12 kHz)
- Amplituda – udává hlasitost zvuku (decibel)

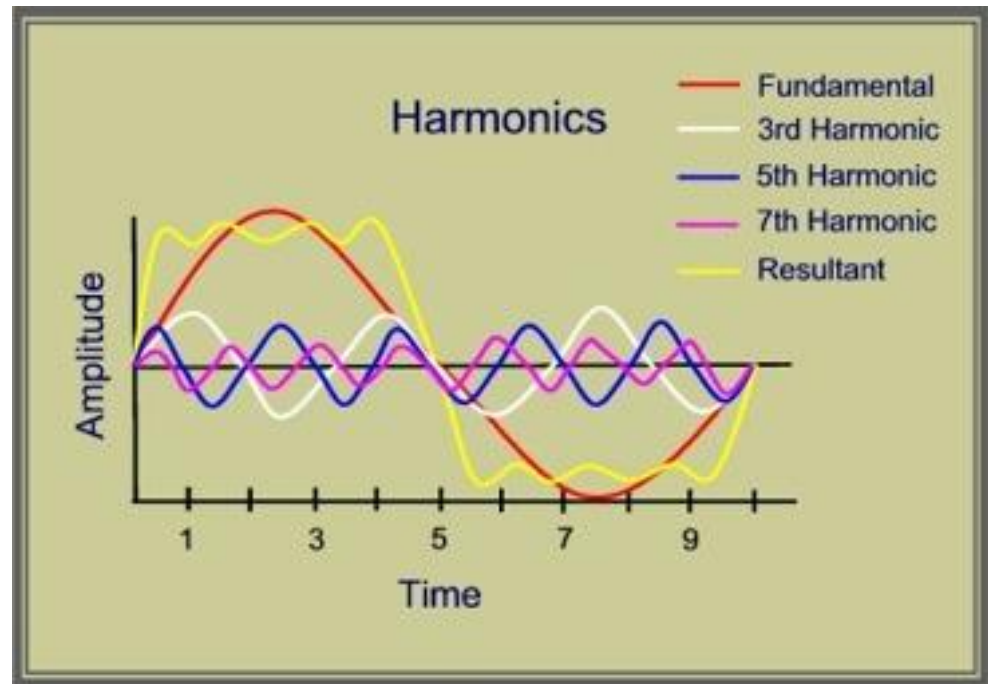


Fyzikální vlastnosti zvuku

- Vlnová délka – vzdálenost mezi dvěma vrcholy vlny (m)
- Šířka pásma – rozdíl mezi nejvyšší a nejnižší frekvencí obsaženou v signálu
- Dopplerův efekt – popisuje změnu frekvence a vlnové délky signálu způsobenou nenulovou vzájemnou rychlostí zdroje a přijímače
 - příklad = siréna na jedoucím autě vydává tón o stále stejné výšce. Pozorovateli se jeví tón sirény auta, které se k němu blíží jako vyšší (vyšší frekvence), než je skutečný tón sirény a naopak tón sirény vzdalujícího se auta se mu jeví jako nižší (nižší frekvence)

Fyzikální vlastnosti zvuku

- Harmonické průběhy – málokterý objekt produkuje zvuk jedné frekvence
- Jediný tón určitého hudebního nástroje je složen ze signálů různých frekvencí => díky tomu snadno rozlišíme jednotlivé nástroje



Digitální zvuk

- Zvuk je analogový
- Pro jeho reprezentaci v PC musí být převeden do digitální podoby
- Mikrofon převádí zvuk na elektrický signál, který je zesílen a filtrován, následně A/D převodníkem převeden do digitální podoby a uložen

X

- Datový tok je posílán na reproduktory skrz A/D převodník, rekonstrukční filtr a zesilovač

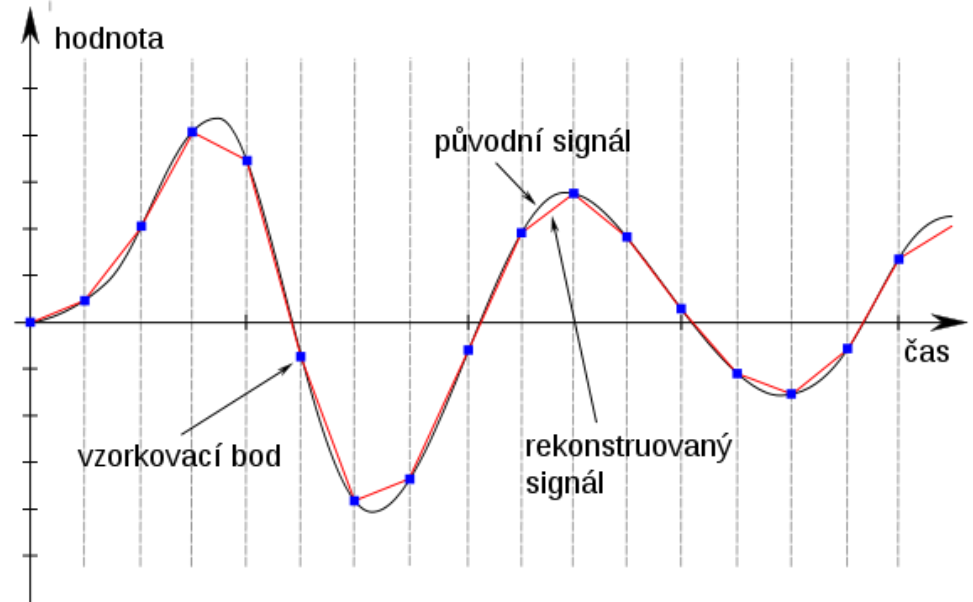
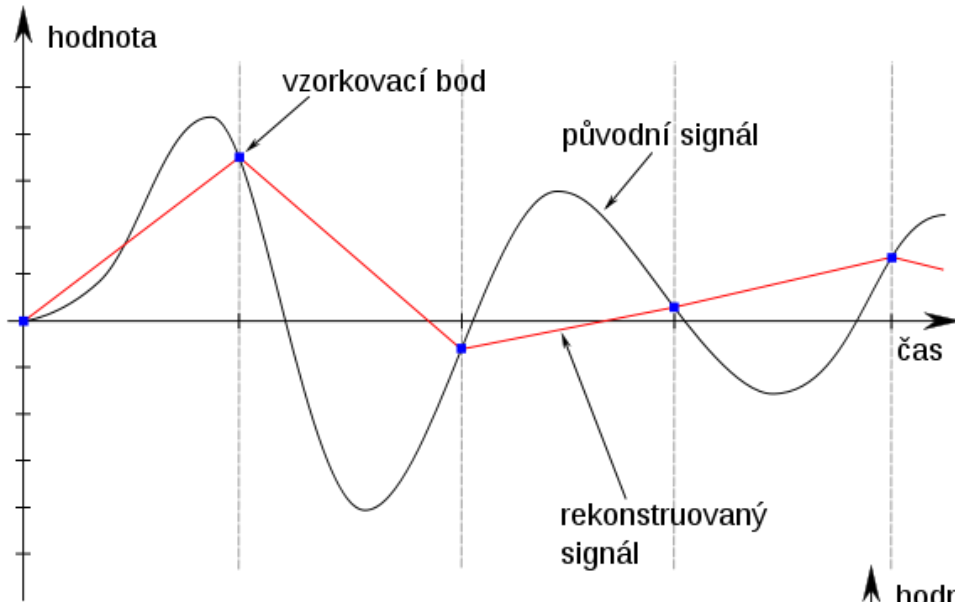
Digitální zvuk

- Zvuk je kontinuální vlnou
- Digitální zvuk vzniká pravidelným vzorkováním této vlny (několik tisíckrát za sekundu)
- Využití Pulzně kódové modulace (PCM) = metoda převodu analogového signálu na digitální
- Princip = pravidelné odečítání hodnoty signálu pomocí A/D převodníku a jejího záznamu v binární podobě
- Každý vzorek zachycuje podobu originální vlny v určitém časovém okamžiku

Digitální zvuk

- Analogie = vznik animace či videa
- Vzorkovací frekvence (sampling rate) = 5 000 – 90 000 vzorků/s
- Vzorkovací frekvence = musí být min. 2x větší než nejvyšší frekvence signálu (aby bylo možno reprezentovat signál korektně), Nyquistův-Shannonův teorém
- Zvuk na CD:
 - vzorkovací frekvence 44.1 kHz => nejvyšší možná korektně zachytitelná frekvence = $44.1/2 = 22\ 050$ Hz (zcela dostačující)

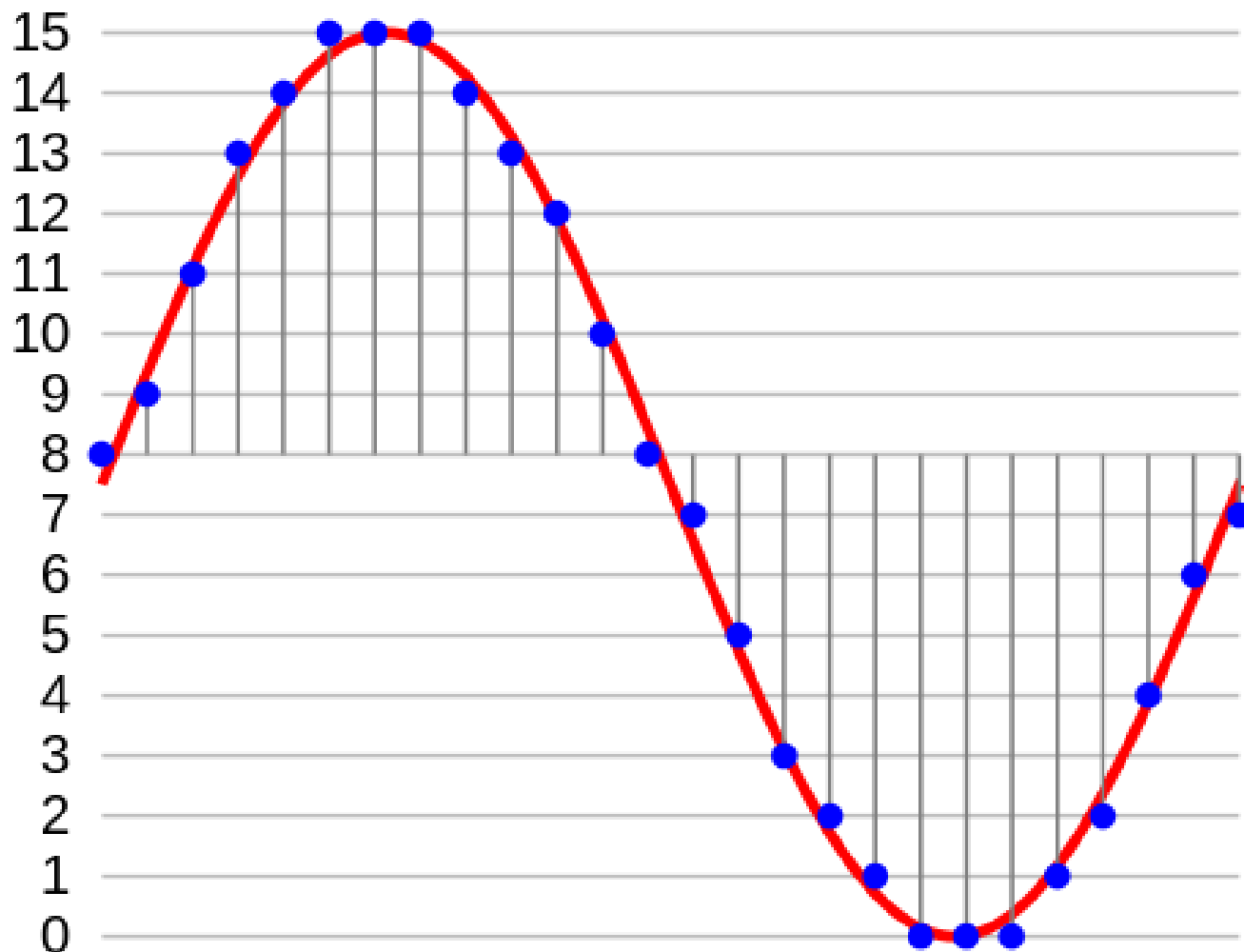
Vzorkovací frekvence a kvalita



Digitální zvuk

- Digitalizace = proces přiřazení **celého** čísla z diskrétní množiny každému vzorku
- 8-bit => 256 hodnot (-128 až 127)
- 16-bit => 65 536 (-32 768 až 32 767)
- Proces digitalizace vnáší do signálu **šum** tím, že musíme zadat celé číslo na určité stupnici (ta nemá natolik jemné členění, abychom celým číslem vyjádřili skutečnou přesnou hodnotu)
- Vyšší bitová hloubka vede k věrnější / přesnější reprezentaci skutečného zvuku

Vliv bitové hloubky digitalizace



Věrnost digitálního zvuku

- Věrnost / kvalita digitálního zvuku je dána kombinací:
 - Vzorkovací frekvence
 - Bitová hloubka digitalizace
- Hi-fi zvuk (high fidelity):
 - 44.1 kHz na kanál (2 kanály = stereo)
 - 16-bit
 - 176.4 kB/s
- Většina aplikací v PC nemá tak vysoké nároky na kvalitu zvuku
- Mono x Stereo

Věrnost digitálního zvuku

Vzorkovací f (kHz)	Bitová hloubka (bit)	Stereo / Mono	Velikost 1minutového souboru (MB)	Poznámka
44	16	Stereo	10.5	CD kvalita
44	16	Mono	5.25	Dobrá kvalita pro komentáře
44	8	Stereo	5.25	Dobrá kvalita pro playback
44	8	Mono	2.6	
22	16	Stereo	5.25	Dobrá, ale ne CD kvalita – FM rádio
22	16	Mono	2.6	OK pro vyprávění
22	8	Stereo	2.6	
22	8	Mono	1.3	AM rádio
11	16	Stereo	2.64	Žádná výhoda z použití sterea
11	16	Mono	1.32	Nejnižší standard pro použití
8	8	Mono	0.47	Telefon

Komprese, kodeky, formáty

Kompresa

- 1 minuta záznamu v CD kvalitě = 10.5 MB/s!
- Historicky klasické techniky bezztrátové komprese:
 - Huffmanovo kódování (slovníková metoda - opakující se sekvence znaků se ukládají kratší sekvencí znaků, kompresní poměr přibližně 1:2)
 - Pokročilejší varianty pracující v několika sekvencích (kompresní poměry cca až 1:3)
 - **Nedostatečné!**
- Aktuálně psychoakustické metody (ztrátové) komprese:
 - Mpeg
 - Ogg
 - ...

Práh slyšení

- sluchový vjem nezávisí pouze na fyzikálních vlastnostech akustické vlny, ale také na vlastnostech sluchového orgánu
- Ten má pro různé kmitočty různou citlivost a dynamický rozsah

Maskování

- Jev, při němž některá část v signálu vjemově překryje jinou, která se tak stává neslyšitelnou
- Při přenosu reálných signálů dochází k maskování neustále a ve velké míře ovlivňuje naše vnímání zvuku
- Prvek signálu, který způsobí překrytí jiného prvku nazýváme maskující, masker
- Prvek překrytý, nazýváme maskovaný, maskee

Kmitočtové maskování

- Maskující a maskovaný prvek se vyskytují v signálu současně
- Silnější komponenta překryje tu slabší
- Překrytá komponenta tak přestane být slyšitelná
- Pokud přehrajeme pouze signál tvořený silnější komponentou, výsledný zvuk nezměníme

Časové maskování

- Maskující prvek předchází nebo následuje maskovaný prvek
- Následuje-li po hlasitém tónu velice podobný tón výrazně nižší hlasitosti, je jeho vnímání potlačeno => dáno automatickou ochranou sluchového ústrojí regulací citlivosti
- V tichém prostředí přenášejí kůstky veškerou akustickou energii z bubínku
- Při silnějším zvuku (větším akustickém tlaku) se do vnitřní části ucha se přenáší jen malá část akustické energie a tím je sluchový orgán chráněn před poškozením
- Tato vlastnost lidského sluchu maskuje tiché tóny následující po hlasitému tónu s podobnou frekvencí a může jí využít algoritmus pro kompresi

Bit rate

- Bit rate = počet kilobitů za sekundu, kterou může komprimovaný soubor se (zvukovým) záznamem využívat
- Čím vyšší bit rate, tím se bude zvukový záznam kvalitou více blížit originálu, ale tím větší bude i velikost jeho souboru
- Konstantní (CBR) – stejný po celou dobu záznamu, snadněji realizovaný
- Proměnlivý (VBR) – v čase proměnlivý (určité části záznamu vyžadují nižší bit rate, než jiné = ticho, jeden nástroj, apod.), obtížnější kódování i dekódování

Bit rate

- Nekomprimovaný zvukový záznam = 1 411 kb/s
- Typické bit rate:
 - 128 kb/s (1:11)
 - 160 kb/s (1:9)
 - 192 kb/s (1:7)

CODEC

- **CO**mpression + **DE**Compression
- Zařízení či software / nástroj provádějící komprimování a **dekomprimování** audia (video) podle určitého algoritmu (metody komprese)
- Kvalita přehrávání zvuku nezávisí jen na použité technice komprese a jejím nastavení, ale také na kvalitě její implementace = kvalitě kodeku
- Nezaměňovat s formátem či kontejnerem!
- Známé kodeky pro formát mp3 = LAME, FhG, Blade, Xing

MPEG-1 = MPEG-2 Audio Layer III = mp3

- Technika ztrátové komprese digitálního audio záznamu + formát pro ukládání zvukového záznamu
- Definovan skupinou MPEG - Motion Pictures Experts Group
- „Založen na nedokonalosti lidského sluchu“
- Ze vstupního signálu odebírá informace, které člověk neslyší, nebo si je neuvědomuje – na základě psychoakustického modelu
- Ponechány jsou pouze složky signály potřebné ke správnému vjemu reprodukováného zvuku
- Velmi dobré výsledky pro hudbu, horší pro mluvené slovo

AAC

- Advanced Coding Audio
- Metoda ztrátové komprese a formát pro ukládání zvukových stop
- Rozvíjí mp3, umožňuje dosahovat vyšší kvality zvuku při totožných bit rate
- ISO standard, součást MPEG-4
- Defaultní audio formát pro:
 - YouTube
 - Zařízení Apple
 - Nintendo
 - Playstation 3

WAV

- Waveform audio file format
- Kontejnerový formát pro ukládání zvuku
- Většinou nekomprimovaný - formát LPCM
- Snadné a rychlé zpracování
- Vytvořen společnostmi Microsoft + IBM
- Používán pro přenos zvukových dat mezi různými systémy, jako pracovní formát při zpracování zvuku, apod.
- Aktuálně je pro tyto účely lepší využívat například formát FLAC

FLAC

- Free Lossless Audio Codec
- Formát pro ukládání zvukových záznamů s využitím bezeztrátové komprese
- Otevřený
- Kompresní poměr obvykle okolo 1:2 (50 – 60 % velikosti původního souboru)
- Velmi rychlá komprese / dekomprese souboru
- Vhodný pro archivaci zvukových záznamů

Ogg (Vorbis Ogg)

- Kontejnerový formát pro multimédia (zvuk, video, titulky, metadata)
- Otevřený
- Využívaný řadou komerčních i nekomerčních zařízení / aplikací
- Podporuje několik typů komprese a formátů audia i videa
- Podporované formáty audia:
 - Ztrátové (Speex, Vorbis, Opus)
 - Bezeztrátové (FLAC)
 - Nekomprimované (PCM)

WMA

- Windows Media Audio
- Kontejnerový zvukový formát pro Windows – proprietární
- Primárně využívá ztrátovou kompresi
- Původně cílen jako náhrada za patentované mp3

Dolby Digital

- Stejnojmenně = AC-3
- Technologie / formát pro ukládání prostorového zvuku (5.1, 7.1, apod.)
- Ztrátová komprese
- Vysoké bit-rate
- Využití pro domácí kina apod. (DVD, Blue-ray, ...)

Zvuk v MMS

- **Zvuk tvořící obsah**

- Komentář
- Zvuky obsažené ve videu, animaci
- Zvuky dávající instrukce = navigace v rámci MSS, apod.
- Hudba = text písně může sdělovat potřebné informace (zdvojit psané informace, apod.)

- **Dokreslující zvuk**

- Hudba = pro vytvoření atmosféry a nálady uživatele
- Ruch, šum na pozadí = pro dokreslení atmosféry (zvuky dopravy z ulice, sportovního utkání, ...)
- Zvukové efekty = např. stisknutí tlačítka navigace, přechod na další stránku, ...

Zvuk v MMS - jak reprezentovat?

- Mluvené slovo – vyžaduje potřebný (profesionální) hlas
- Zvuková stopa – nahrání na mikrofon
- MIDI – vygenerování v PC