

2. Akustika a vizualizácia reči pomocou PRAAT-u

2.1. Úvod

Cieľom tejto kapitoly je predstaviť prístupnou formou základy akustiky reči a možnosti softvéru PRAAT pri vizualizácii akustických vlastností reči. V tomto prehľade sa sústredíme hlavne na tie oblasti, ktoré sú nevyhnutné pre praktickú prácu s nahrávkami reči, ktorú popisujeme v Kapitole 3. Keďže ide o prepojenie teoretických fyzikálnych vlastností s praktickými ukážkami týchto vlastností pomocou PRAAT-u, nevyhnutne tak budeme predstavovať dva okruhy vedomostí a zručností, ktoré predpokladáme sú pre čitateľa do určitej miery nové. Pozitívom tohto praktického prístupu je podľa nás možnosť vyskúšať si vizualizáciu prezentovaných javov a tým pomocou zapojenia viacerých kognitívnych systémov tieto javy hlbšie pochopiť. Taktiež môže čitateľ interaktívnou formou získať základnú orientáciu v softvérovom prostredí PRAAT-u, ktorá bude prospešná a potrebná pri ďalších kapitolách. Potenciálnym problémom tohto prístupu je tá skutočnosť, že je nemožné postupovať absolútne súbežne v teoretickej a praktickej oblasti, ale budeme sa snažiť udržiavať vzdialenosť medzi oboma oblasťami na minime. V prílohe A uvádzame manuál pre prácu s PRAAT-om, ktorý prehľadne a na jednom mieste sumarizuje postupy podrobnejšie prezentované v texte kapitoly.

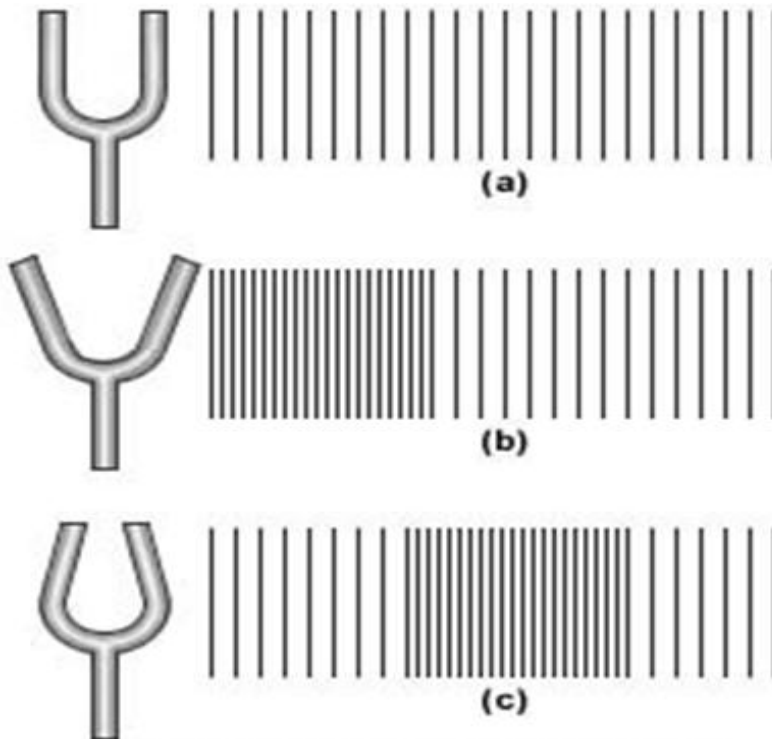
2.2. Reč a jej fyzikálne vlastnosti

Produkcia a percepcia ľudskej reči je komplexný mechanizmus, ktorý vytvárajú spolu fyziologická činnosť artikulačných a percepčných orgánov na jednej strane a kognitívny systém vzťahov medzi charakteristikami týchto fyziologických činností a významovými rozdielmi v rôznych rovinách jazykovej komunikácie na strane druhej. Napríklad po formálnej stránke môžeme skúmať rozdiel medzi krátkou a dlhou samohláskou /a/. Zistíme, že krátke /a/ je nielen kratšie ako dlhé /á/, ale aj to, že krátke /a/ je viac centralizované, teda vyslovuje sa s menším otvorom čeluste a vyššou polohou jazyka ako je tomu pri dlhom /á/ (Beňuš a Mády 2010). Po funkčnej stránke však tieto dva rozdiely nemajú rovnaké postavenie, pretože kontrast medzi dlhou a krátkou samohláskou v slovenčine signalizuje významové rozdiely (aj keď funkčné zaťaženie tohto kontrastu nie je veľké) ako napr. pri slovách *látka-latka*, alebo *putá-púta*. Naproti tomu rozdiel medzi viac periférnou kvalitou a viac centralizovanou kvalitou samohlások v slovenčine takýto významový kontrast nesignalizuje, a je preto funkčne na inej úrovni, ako kontrast dĺžkový.

Túto časť začneme popisom a vizualizáciou čisto fyzikálnych vlastností zvuku, ktoré sú nezávislé od toho, ako bol zvuk vyprodukovaný alebo vnímaný. Inými slovami,

budeme sa venovať charakteristikám zvukového signálu medzi (a vynímajúc) jeho producentom a prijímateľom. Na základe týchto poznatkov potom prejdeme k tomu, ako tieto fyzikálne vlastnosti súvisia s rečou, a ako nám ich vizualizácia pomôže lepšie analyzovať úlohu fonetických charakteristík v celej škále výrazových prostriedkov jazyka.

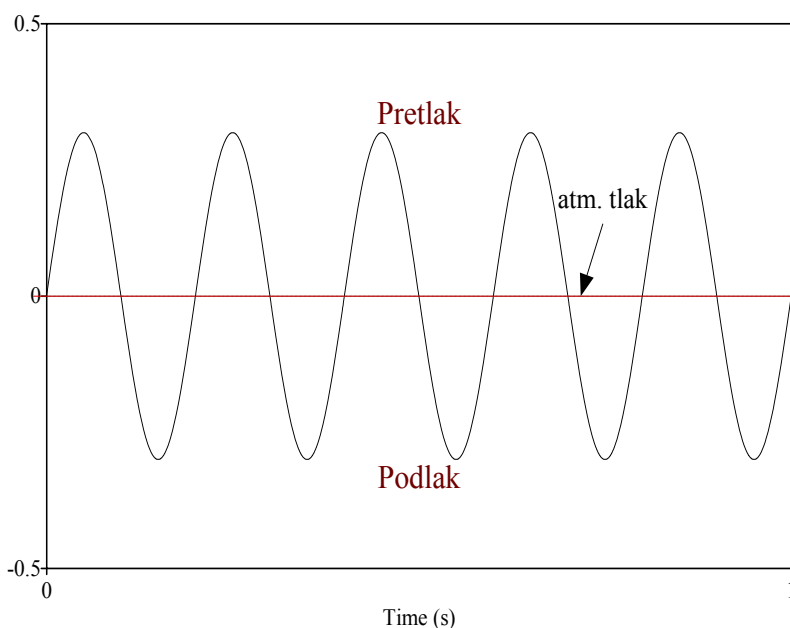
Zvuk z fyzikálneho hľadiska je kmitavý pohyb častíc, ktorý spôsobuje zmeny tlaku. Najčastejšie hovoríme o pohybe vzduchových molekúl a následných zmenách atmosférického tlaku, ale zvuk môže samozrejme predstavovať aj kmitanie molekúl vody alebo inej elastickej látky, v ktorej majú molekuly možnosť pohybu. Pri zvuku sa tak v atmosférickom tlaku vytvárajú miesta zvýšeného a zníženého tlaku. Môžeme si to ilustrovať na príklade ladičky na Obrázku (1).



Obrázok (1) Zmena tlaku spôsobená kmitaním ladičky
(<http://home.att.net/~cat4a/wave-II.htm>)

Na hornom paneli (a) je ladička v pokoji a zvislé čiary znázorňujúce molekuly vzduchu sú od seba rovnako vzdialené pri normálnom atmosférickom tlaku. Na

strednom paneli (b) sa nárazom ladičky o tvrdý predmet a vďaka jej pružnosti jej konce vychýlia, čím vytlačia molekuly vzduchu a vytvorí sa lokálne zhustenie (pretlak). Pri spätnom pohybe ladičky znázornenom na spodnom paneli (c) je hustota molekúl nižšia a vzniká podtlak, keďže molekuly, ktoré tam boli pôvodne boli vytlačené vo fáze (b) a nijaké nové molekuly neboli dodané. Porovnaním fázy (b) a (c) taktiež vidíme, že lokálna oblasť pretlaku sa pohybuje v čase smerom od zdroja a vzniká tak zvuková vlna. Následné vychýlenia ladičky spôsobia nové miesta pretlaku a podtlaku, ktoré sa budú šíriť smerom od zdroja. Tento jav je podobný kruhovým vlnám na pokojnej vodnej hladine, ak do nej hodíme kameň. Je dôležité si uvedomiť, že sa šíri zvuková vlna, teda striedajúce sa miesta zvýšeného a zníženého tlaku, ale jednotlivé molekuly vzduchu iba kmitajú na svojom mieste.



Obrázok (2) Znázornenie zmien tlaku sínusoidou.

Vhodným zobrazením zvukovej vlny je graf znázorňujúci zmeny tlaku (os y) v čase (os x). Na Obrázku (2) vidíme striedanie zvýšenej a zníženej hustoty molekúl vzduchu, teda pretlaku a podtlaku, na stabilnom mieste v čase plynúcom zľava doprava. V tomto konkrétnom prípade sa tlak zvýši a zníži päťkrát za jednu sekundu, čiže frekvencia tejto takejto zvukovej vlny je 5 Hz.

Ľudské ucho je schopné vnímať frekvencie zhruba od 20 do 20 000 Hz, ale v reči sa využíva len dolná polovica tohto intervalu. V ľudskej reči sa takéto zmeny tlaku vytvárajú napríklad otváraním a zatváraním hlasivkovej štrbiny niekoľko stokrát za sekundu, a po prenose vzduchom takejto zvukovej vlny sa dostávajú na citlivé membrány sluchového ústrojenstva, za pomoci ktorých ich ľudská myseľ vníma ako reč.

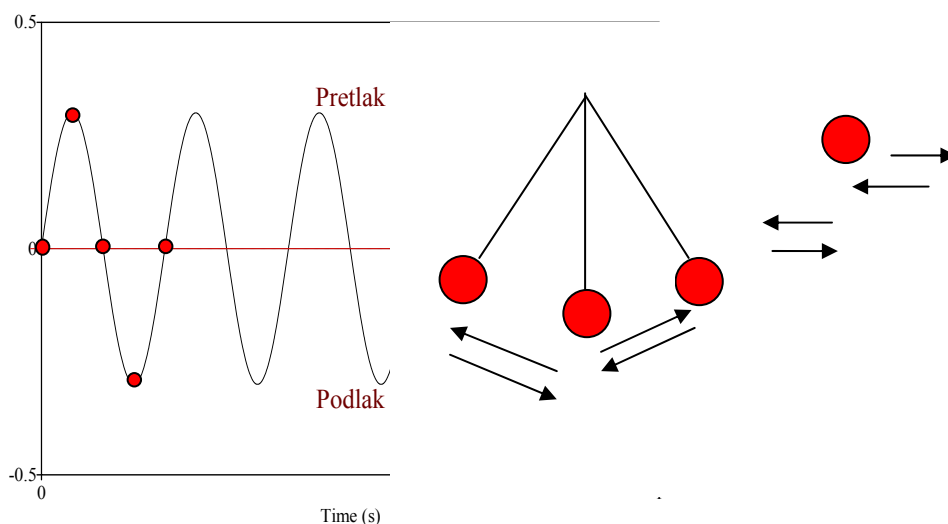
Na tomto mieste je vhodné začať sa oboznamovať so softvérom na vizualizáciu, fonetickú analýzu a manipuláciu reči PRAAT, aby sme si mohli názorne vizuálne a sluchovo predstaviť niektoré základné fyzikálne charakteristiky zvukových vln. PRAAT je voľne dostupný program, ktorý sa dá stiahnuť z web-stránky www.praat.org pre najbežnejšie operačné systémy (Windows, Mac, Linux, Unix). Po kliknutí na túto stránku kliknite v ľavom hornom rohu na operačný systém, ktorý je nainštalovaný na počítači, na ktorom budete s PRAAT-om pracovať (Windows, Mac alebo Linux). Postupujte podľa jednoduchého návodu na tejto strane. U novších verzií operačných systémov stačí uložiť .exe súbor najnovšej verzie PRAAT-u (napr.praat5013_winsit.exe z 18. 3. 2008) na pevný disk Vášho počítača. Po uložení takéhoto súboru naň dvakrát kliknite, a potom kliknite OK aby sa PRAAT nainštaloval na plochu Vášho počítača. V prípade, že máte na počítači, na ktorom budete PRAAT používať, jednu zo starších verzií Windows, je možné stiahnuť si staršiu verziu PRAAT-u. PRAAT je neustále zlepšovaný, preto je dobré pravidelne si sťahovať najnovšiu verziu programu. Po úspešnom nainštalovaní kliknite na ikonu PRAAT-u, ktorá by sa mala objaviť na ploche (ružové ústa s jazykom). V prípade, že nemáte prístup na internet, alebo stránka PRAAT-u je z akýchkoľvek príčin nedostupná, v priečinku Praat na DVD, ktoré je súčasťou tejto monografie, sa nachádza verzia softvéru pre OS Windows z 4. Augusta 2010. Pri inštalovaní tejto verzie PRAAT-u postupujte rovnako, teda dvakrát kliknite na .exe súbor, a potom kliknite OK aby sa PRAAT nainštaloval na plochu Vášho počítača.

Po spustení programu (kliknúť dvakrát na jeho ikonu na ploche) sa Vám otvoria dve okná: *Praat Objects*, ktoré sa používa na základnú analýzu, a *Praat Picture*, ktoré sa využíva na vytváranie obrázkov zvukového signálu, alebo anotácie. V okne *Praat Objects* je na lište päť možností: *Praat, New, Read, Write, Help*.

Teraz použijeme PRAAT na vytvorenie prvého zvukového súboru. V priečinku "Programy_pre_uzivatela" na DVD, ktoré je súčasťou tejto monografie, sa nachádzajú rôzne jednoduché programy, ktoré automatizujú niektoré úkony v PRAAT-e. V okne *Praat objects* kliknite na *Praat* na hornej lište a *Open Praat Script* a navigujte do tohto priečinku na DVD a otvorte súbor *zvuk_sinusoida.praat*. Otvorí sa Vám nové okno s týmto niekoľkoriadkovým programom, ktorý spustíte z lišty tohto okna cez Run → Run

(Ctrl-R). Po spustení programu sa otvorí okno, kde môžete zadať základné charakteristiky pre novovytvorený zvuk: jeho frekvenciu v Hertzoch, dĺžku v sekundách, a parameter hlasitosti od 0 po 1 (predvolené hodnoty sú 440, 1, a 0.5). Ponechajte predvolené hodnoty, kliknete na *OK*, a ak máte správne nainštalovaný zvukový výstup, PRAAT vyprodukuje referenčný tón danej frekvencie a dĺžky. Všimnite si, že v okne *Praat Objects* sa objavil nový súbor nazvaný *sine440*, ktorý obsahuje Vami vytvorený zvuk. Ak by ste chceli tento súbor uchovať, treba naň kliknúť, aby sa vysvietil na modro, a potom z lišty: *Write* → *Write to wav file...* a navigujte tam, kde si súbor chcete uložiť. Ak by ste si ho chceli znovu vypočítať, len kliknite na tlačidlo *Play* na pravej strane okna *Praat Objects*. Pohrajte sa s týmto programom opätovným spúšťaním súboru *zvuk_sinusoidea.praat* cez CTRL-R a nastavením rôznych frekvencií.

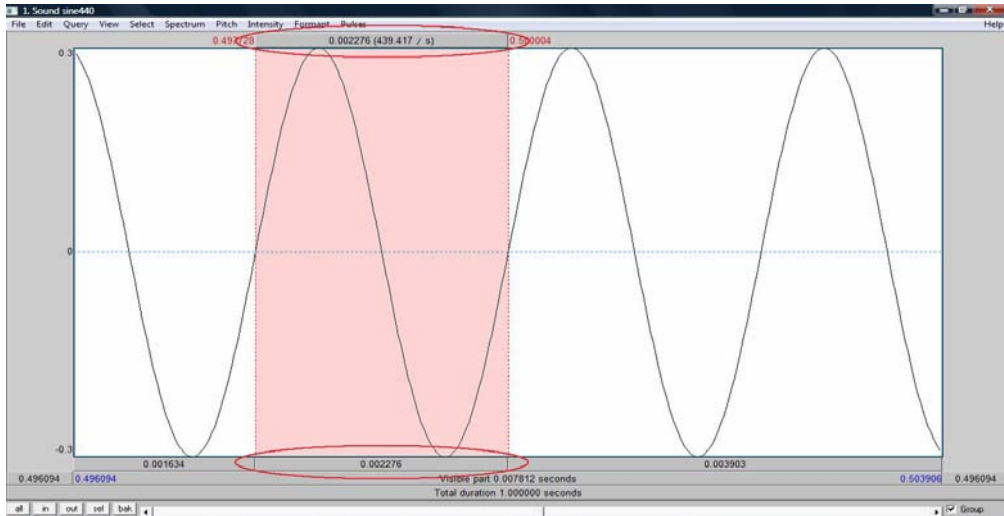
Teraz sa bližšie pozrime na niektoré charakteristiky takéhoto zvuku. Vytvorte zvuk s frekvenciou 5 Hz, dĺžkou 1 sekunda, a hlasitostným parametrom 0,3. V tomto prípade nebudete nič počuť, lebo ľudské ucho neregistruje frekvencie nižšie ako 20 Hz. V okne *Praat Objects* sa Vám objaví teda súbor *sine5*, kliknite naň aby sa vysvietil na modro, a z možnosti na pravej strane okna kliknite na *Edit*. Otvorí sa Vám nové okno s podobnou sínusoidou, akú sme videli na Obrázku 2. Jednou z možností interpretácie takejto sínusoidy je predstaviť si ju ako pohyb jednej molekuly vzduchu ako keby bola zavesená na kyvadle (alebo upevnená na pružine), ako je to ilustrované na Obrázku (3).



Obrázok (3) Sínusoida ako kyvadlový pohyb jednej molekuly vzduchu

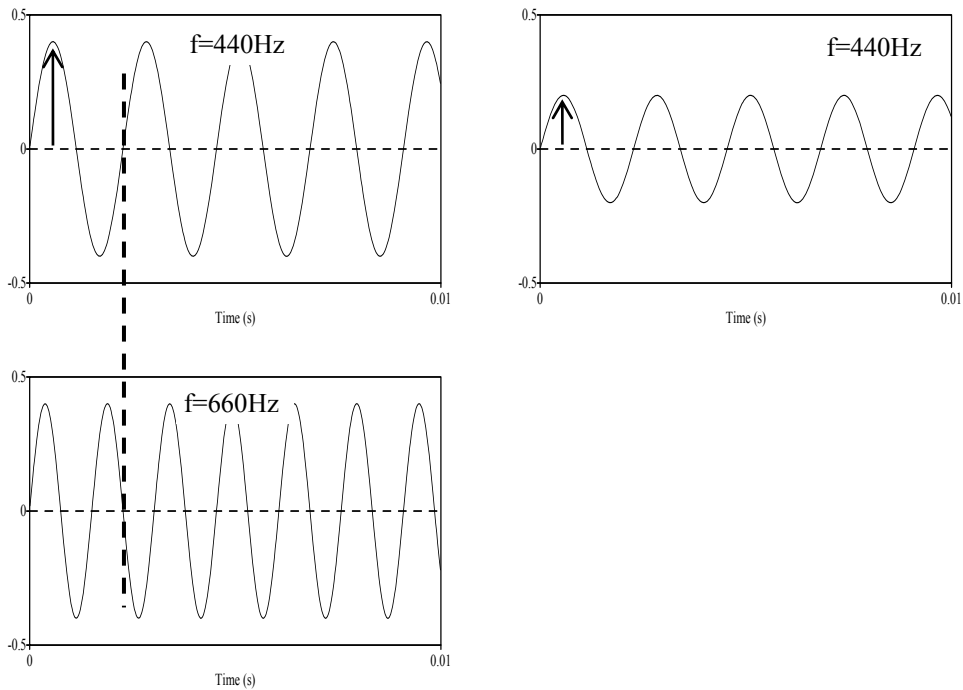
Takéto kyvadlo predstavujúce molekulu sa po úvodnom vychýlení z rovnovážnej polohy spôsobenom vonkajším vplyvom, v našom prípade kmitajúcim telesom alebo okolitými molekulami, vráti naspäť do rovnovážnej polohy a v dôsledku zotrvačnej energie sa vychýli na opačnú stranu a naspäť do rovnovážnej polohy. Tento pohyb predstavuje jeden cyklus, a na sínusoide jednu periódu.

Takže zvuková vlna *sine5*, ktorú sme si vytvorili v PRAAT-e je typickým príkladom periodickej vlny, keďže znázorňuje periodický pohyb opakujúci sa v pravidelných intervaloch. Naša vlna má 5 cyklov, ktoré prebehnú za 1 sekundu, teda frekvenciu 5 Hz, a trvanie jedného cyklu je $1/5 = 0.2$ sekúnd. Frekvencia je vždy počet cyklov za jednu sekundu. Ako dlho trvá cyklus v *sine440*, teda v prvej zvukovej vlne, ktorú sme vytvorili? Jedným spôsobom ako to zistiť je jednoducho odmerať tento interval pomocou PRAAT-u. Kliknite na *sine440* v okne *Praat Objects* a z možností na pravej strane okna kliknite na *Edit*. Otvorí sa Vám nové okno s týmto zvukom. Keďže cyklov je v tomto prípade 440 za sekundu, musíme zobrazený interval zmenšiť a tým si zvukovú vlnu priblížiť podobnou ako pri mikroskope alebo fotografovaní. Na tieto operácie slúžia v PRAAT-e tlačidlá vľavo dole editovacieho okna. *All* zobrazí sa celý súbor, *In* priblíži (CTRL-I), *Out* vzdiali (CTRL-O), *Sel* zobrazí označenú časť sa v celom okne (CTRL-N), a *Bak* vráti zobrazenie na predošlú úroveň. Všetky tieto možnosti sa nachádzajú aj na hlavnej lište pod možnosťou *View*. Takže na približné odmeranie jedného cyklu v *sine440* si potrebujeme niekoľkokrát signál priblížiť pomocou *In* (CTRL-I), až kým jasne nevidíme jednotlivé cykly sínusoidy (napr. sedemkrát). Keď teraz myšou označíme úsek od začiatku do konca sínusoidy tak ako je to na Obrázku (4), PRAAT zobrazí dĺžku tohto úseku ako 0,002276 sekúnd. Matematicky je dĺžka jedného cyklu presne $1/440 = 0,0022727$.



Obrázok (4) Približné meranie dĺžky jedného cyklu v PRAAT-e.

Amplitúda a frekvencia zvukovej vlny sú jej základnými charakteristikami. Je dôležité si uvedomiť, že tieto dve vlastnosti sú nezávislé od seba. Amplitúda vyjadruje, ako veľmi sa molekula vzduchu pri kmitaní vychýli zo svojej rovnovážnej polohy. Je teda vyjadrením tlaku, ktorý na ňu pôsobí, a tým hlasitosti zvuku. Frekvencia vyjadruje koľkokrát za sekundu sa toto vychýlenie opakuje a charakterizuje výšku zvuku. Nezávislosť týchto dvoch veličín je ilustrovaná na Obrázku (5). Zvuková vlna v ľavom hornom paneli má frekvenciu 440 Hz a maximálnu amplitúdu 0,4. Zvuková vlna pod ňou, v ľavom dolnom paneli, má frekvenciu 660 Hz a maximálnu amplitúdu 0,4. Takže tieto dve zvukové vlny majú rovnakú amplitúdu, ale rozdielnu frekvenciu. Vidíme, že jeden cyklus 440 hertzovej vlny zodpovedá jeden a pol cyklu 660 hertzovej vlny. Porovnanie sínusoid v ľavom a pravom hornom paneli zase ukazuje prípad, keď sa zhoduje frekvencia ale amplitúda je rozdielna. Samozrejme, že všetky tri zvukové vlny si môžeme vytvoriť v PRAAT-e pomocou už známeho programu *zvuk_sinusoida.praat* a porovnať si ich aj sluchovo.

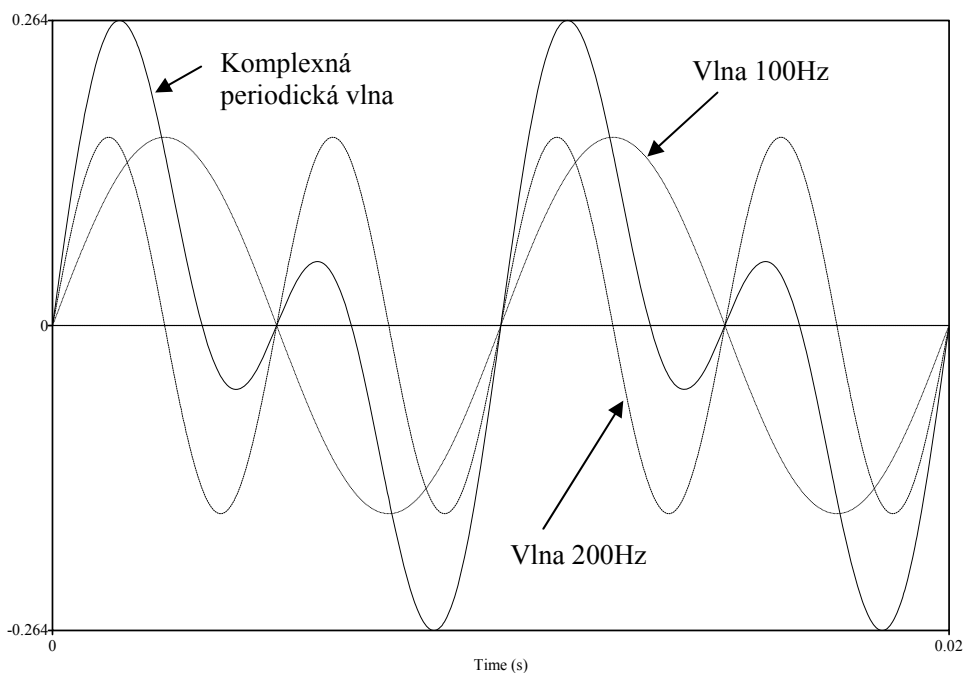


Obrázok (5) Nezávislosť frekvencie a amplitúdy zvukovej vlny

Zvuky, ktorými sme sa doteraz zaoberali sú takzvané referenčné zvuky, ktoré sú vyjadrené jednoduchou sínusoidou. Ako ste však mohli pozorovať pri posluchu týchto zvukov v PRAAT-e, veľmi sa na ľudskú reč nepodobajú. Reč sa skladá hlavne z komplexných periodických zvukov, neperiodických zvukov, a ich kombinácií. Najskôr sa pozrime na komplexné periodické zvuky. Tieto zvuky vznikajú kombináciou viacerých jednoduchých periodických zvukov. Hore vyššie sme už popísali, ako si pomocou programu *zvuk_sinusoidea.praat* vytvoríme zvuk s frekvenciou 440 Hz. Použite znovu tento skript a vytvorte dva zvuky¹ jeden s frekvenciou 100Hz (*sine100*) a druhý s frekvenciou 200Hz (*sine200*). Myšou označte obidva súbory v okne *Praat Objects* a skombinujte ich do stereo zvuku (*Combine sounds* → *Combine to stereo*), čím sa v okne Praat Objects objaví zvuk s názvom *sine100_sine200*, a potom ho konvertujte na mono zvuk (*Convert* → *Convert to mono*), čím vytvoríte zvuk *sine100_sine200_mono*. Týmto postupom sme vlastne spočítali prvú a druhú vlnu.¹ Kliknite na tento nový

¹ Nie je to úplne presne, pretože Praat pri týchto dvoch krokoch prispôsobí amplitúdu, ale pre naše ilustračné účely to postačuje.

zvuk, prehrajte pomocou *Play*, a potom si ho pozrite stlačením *Edit*. Pomocou tlačidiel v ľavom dolnom rohu editovacieho okna (alebo skratiek Ctrl-I a Ctrl-O) si vlnu priblížite a zistíte, či je táto vlna periodická, a ak áno, aká je dĺžka jedného cyklu. Zistíte, že aj napriek komplexnejšiemu vyzerajúcemu priebehu je táto vlna periodická a jej cyklus je dlhý 0.01 s, čiže jej základná frekvencia, často označovaná F0, je 100 Hz. Pri komplexných periodických vlnách je základná frekvencia F0 vždy zodpovedajúca frekvencii najnižšieho komponentu, v tomto prípade teda 100 Hz keďže sme kombinovali 100 a 200 Hz vlny. Ilustračné znázornenie komplexnej vlny s F0 100Hz, ktorá vznikne pridaním 100Hz a 200Hz vlny je uvedené na Obrázku (6). Každá komplexná vlna sa skladá z čiastkových vln, ktoré majú frekvenciu vyjadriteľnú ako celočíselný násobok základnej frekvencie F0.



Obrázok (6) Komplexná periodická vlna, ktorá vznikne pridaním 100Hz vlny a 200Hz vlny.

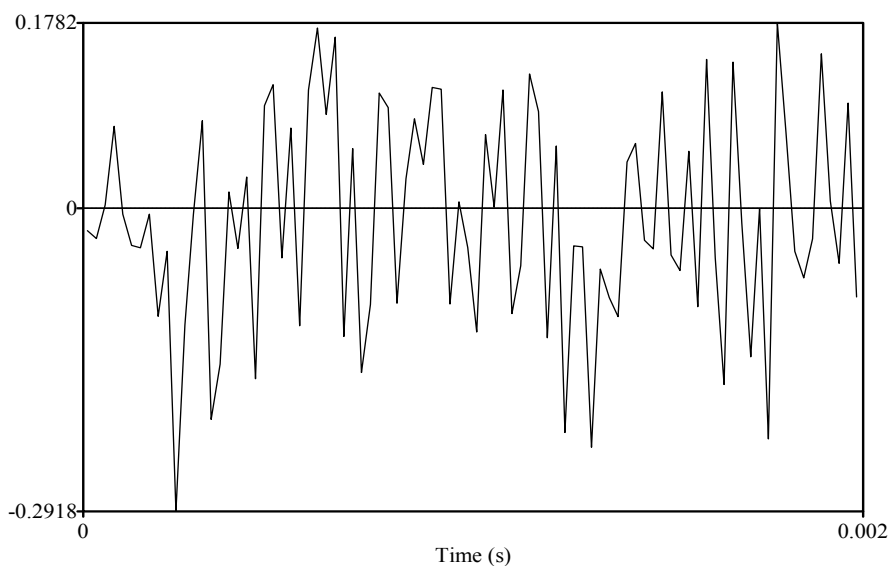
Urobme teraz malý experiment a vytvoríme 300 Hz vlnu a skombinujeme ju s pôvodnou 200Hz vlnou. Ak ste doteraz postupovali podľa návodu, všetko by malo byť jasné, a ak niečo nefunguje, prezrite si ešte raz predošlé strany tejto podkapitoly. V prípade, že potrebujete označiť súbory, ktoré nenasledujú priamo za sebou v okne Praat

Objects, kliknite na prvý súbor myšou, a na každý ďalší súbor kliknite spolu so súčasným stlačením CTRL tlačidla. Aká je základná frekvencia novej vlny? Podľa doterajších skúseností by to malo byť 200 Hz, teda frekvencia najnižšieho komponentu. Ale nie je tomu tak. Kombináciou týchto dvoch vln dostaneme komplexnú periodickú vlnu, ktorej základná frekvencia je opäť 100Hz (pretože jeden cyklus trvá 0.01s). Je to preto, lebo obidve pôvodné vlny sa dajú vyjadriť ako násobok celým číslom tejto základnej frekvencie ($2 \cdot 100$, respektíve $3 \cdot 100$). Keďže každá komplexná periodická vlna je vlastne súčtom čiastkových vln v každom časovom bode (hodnote na osi x), dá sa teda spätným analytickým postupom vyjadriť ako súbor jednoduchých vln, z ktorých sa skladá. Používa sa pri tom viacero matematických metód, z ktorých najznámejšia je Fourierova transformácia, ktorá rozkladá komplexnú periodickú vlnu na súbor sínusových a kosínusových vln. V ľudskej reči sa ako komplexné periodické vlny prejavujú hlavne samohlásky a aj niektoré sonórne spoluhlásky.

Druhou hlavnou skupinou zvukov, ktoré sa využívajú v reči sú zvuky vyjadrené aperiodickými vlnami. V týchto vlnách sa nedá identifikovať žiadny opakujúci sa cyklus, sú preto nepravidelné a majú čiastkové frekvencie, ktoré sa nedajú vyjadriť ako celočíselné násobky základnej frekvencie. Príkladom takéhoto zvuku je napríklad “biely šum“, ktorý má rovnako hlasné čiastkové frekvencie v celej frekvenčnej škále. Keďže Praat obsahuje niektoré predvolené funkcie, takáto vlna sa dá jednoducho vytvoriť nasledovným postupom:

1. V okne Praat Objects kliknite *New* → *Sound* → *Create sound from formula...*
2. V okne, ktoré sa objaví nazvite túto vlnu *Noise* a ponechajte ostatné predvolené hodnoty pre kanál, začiatočný a koncový čas a vzorkovaciu frekvenciu
3. Ako “formula“ vložte: *randomGauss(0,0.1)*
4. Kliknite *OK* a v okne Praat Objects sa objaví nový objekt *Sound Noise*

Opäť si tento zvuk prehrajte pomocou *Play* a prezrite pomocou *Edit*. Priblížením si overte, že takáto vlna je skutočne neperiodická. Na Obrázku (7) sú znázornené dve milisekundy takejto vlny a vidíme, že skutočne sa nedá identifikovať žiadny pravidelne sa opakujúci cyklus.



Obrázok (7) Komplexná aperiodická vlna.

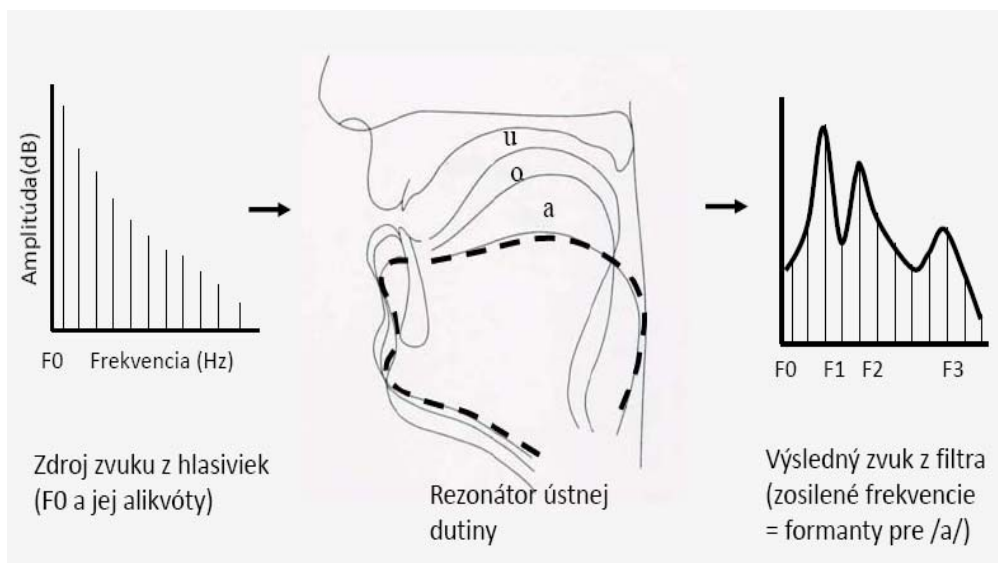
V ľudskej reči sa takýmito vlnami prejavujú hlavne tzv. trené zvuky (frikatívy) ako napríklad /s/ alebo /f/, a tiež krátka fáza otvárania tzv. záverových spoluhlások ako napríklad /p/ alebo /k/.

K dosiahnutiu základného porozumenia akustických vlastností reči nám chýbajú už len dva kroky. Prvým je uvedenie si charakteristiky zdrojového zvuku, ktorý vytvárajú hlasivky. Kmitanie hlasiviek je aerodynamický proces využívajúci tzv. Bernouliho efekt, ktorý zhruba vyjadruje nepriamu úmernosť medzi rýchlosťou pohybu molekúl vzduchu a atmosférickým tlakom: čím rýchlejšie sa molekuly pohybujú, tým je tlak v mieste pohybu menší. V hrtane to funguje zjednodušene tak, že hrtanové svaly regulujú pozíciu hrtanových chrupaviek, na ktorých je pripevnené tkanivo hlasiviek. Svaly tak môžu meniť vzdialenosť hlasiviek od seba a taktiež ich napätosť. Pri určitom nastavení sú teda hlasivky zatvorené a pripravené na kmitanie. Sila tlaku vzduchu, ktorý z pľúc cez priedušnicu pôsobí na zatvorenú hlasivkovú štrbinu, v určitom momente presiahne silu, ktorá udržuje hlasivky zatvorené, a hlasivková štrbina sa otvorí. V tejto štrbine sa však molekuly vzduchu pohybujú veľkou rýchlosťou, lebo rozdiel tlaku, ktorý bol pred otvorením hlasiviek „pod“ nimi a v ústnej dutine „nad“ nimi je veľmi veľký. Bernouliho efekt spôsobí, že tlak vzduchu medzi hlasivkami prudko klesne z dôvodu veľmi rýchleho pohybu vzduchových molekúl. Keďže tento medzihlasivkový tlak je

jedinou silou, ktorá drží hlasivky otvorené, a hlasivkové svalstvo pôsobí silou tak, aby boli hlasivky zatvorené (tak ako pred tým, ako ich tlak vzduchu otvoril), prudké zníženie medzihlasivkového tlaku spôsobí opätovné zatvorenie hlasiviek. Tento cyklus otvárania a zatvárania hlasiviek sa opakuje zvyčajne 80-400 krát za sekundu a určuje základnú frekvenciu reči.

Tým, že hlasivky sú pružné kmitajúce teleso, nevytvárajú referenčné, t.j. jednoduché periodické, zvuky podobné ladičke, ale zvuky zložené z viacerých frekvencií. Tieto alikvotné frekvencie sú vždy celočíselným násobkom základnej frekvencie. Je to podobné, ako napríklad pri strune na husliach alebo gitare. Mechanickým rozozvučením struny brnknutím alebo trením sláčika začne struna kmitať ako jeden celok, čo vytvorí základnú frekvenciu v závislosti od jej dĺžky a napnutia, ale takisto kmitá polovica struny, potom polovica tej polovice, atď. Týmto spôsobom znejú popri najhlasnejšej základnej frekvencii aj alikvotné frekvencie s menšou amplitúdou (teda hlasitosťou). Čiže napríklad tón a1 na husliach má základnú frekvenciu 440Hz, ale popri tejto frekvencii je prítomná aj frekvencia 880Hz, 1320Hz, 1760Hz, atď. Tieto vyššie frekvencie potom dotvárajú farbu zvuku a pri hlasivkách závisia od plasticity hlasivkového tkaniva a pravidelnosti otváracie a zatváracie fázy kmitania.

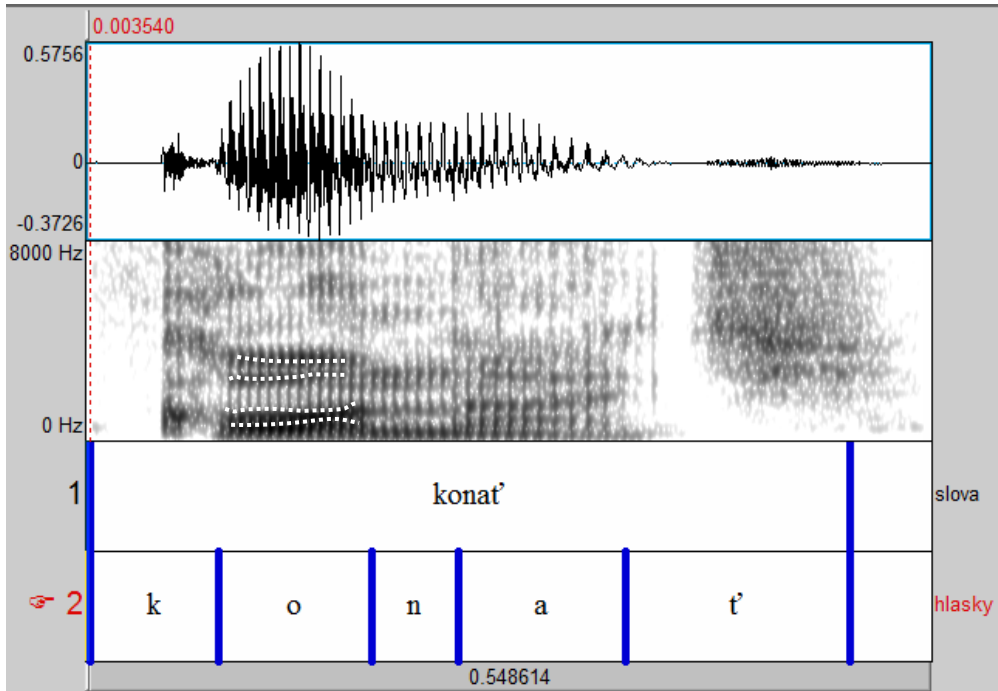
Takže v prvom kroku sme si utvrdili to, že zdrojom zvuku pri ľudskej reči sú zvuky obsahujúce základnú, ale aj alikvotné frekvencie. Druhým krokom je uvedenie si úlohy filtra, alebo rezonátora. Znovu si pomôžeme hudobnými nástrojmi. Na to, aby sa zvuk mohol šíriť, potrebuje byť zosilnený, na čo slúži napríklad pri strunových nástrojoch telo nástroja. Je to dutina rôzneho objemu a zvyčajne nepravidelných tvarov, ktorá dokáže selektívne zosilniť niektoré frekvencie, a naopak niektoré utlmiť. Predstavte si, že huslista zahrá notu a1 na husliach a potom čelista zahrá tú istú notu na violončele. Pravdepodobne budete počuť rozdiel vo farbe zvuku. Tento rozdiel je spôsobený tým, že väčší rezonátor violončela lepšie zosilňuje nižšie z alikvotných frekvencií, a naproti tomu menšie telo huslí zase rozozvučí tie vyššie frekvencie. Na podobnom princípe funguje aj reč. Veľkosť a hlavne tvar ústnej dutiny (spolu s menej dôležitou nosovou dutinou) reguluje to, ktoré alikvotné frekvencie budú zosilnené a ktoré budú utlmené. Tie zosilnené frekvencie sa typicky nazývajú formanty, od najnižšieho F1, postupne F2, F3, atď. Takže zmeny tvaru ústnej dutiny hlavne pôsobením jazyka, čeluste a pier určujú výšku jednotlivých formantov, čo ľudské ucho vníma ako rozdiely v kvalite sonórnych hlások, hlavne samohlások.



Obrázok (8) Ilustratívna ukážka tzv. Source-filter teórie produkcie reči

Obrázok (8) sumarizuje posledné dva kroky, ktoré sme teraz popísali. Vľavo je graf frekvencie (os x) a amplitúdy (os y), ktoré sa nazývajú spektrum. V strede je znázornený filter ústnej dutiny a vpravo výsledné spektrum s približnými formantovými frekvenciami pre samohlásku /a/.

Keďže spektrum na Obrázku 8 ukazuje frekvenciu a amplitúdu len v jednom časovom bode, nie je veľmi vhodné na zobrazovanie reči, ktorá sa prirodzene odohráva v určitom časovom intervale. Preto je bežnejšie pre účely zobrazovania reči využívať spektrogram, ktorý zobrazuje tri dimenzie: čas (os x), frekvencia (os y) a amplitúda, t.j. hlasitosť vyjadrená „tmavosťou“ každého bodu v xy priestore. Oscilogram zvukovej vlny spolu so spektrogramom zobrazujúcim slovo *konat'* je na Obrázku (9). Najlepšie je opäť otvoriť si tento zvukový súbor v PRAAT-e a vyskúšať si tak „čítanie“ spektrogramu.



Obrázok (9) Oscilogram a spektrogram slova "konat".

Postupujte už známym sledom krokov (*Praat Objects* → *Read* → *Read from file...*) a navigujte do priečinku na DVD "Prikklady", kde nájdete súbor *konat.wav* (podľa nastavenia zobrazovania súborov priečinku na Vašom počítači sa môže súbor objaviť ako *konat.wav* alebo iba *konat*). Po načítaní PRAAT-om sa tento súbor objaví v okne *Praat Objects* ako *Sound konat*. Kliknite *Edit*. V novom okne, ktoré sa otvorí nastavte, ktoré grafy má Praat zobrazit': *View* → *Show analyses...* a označte iba prvú možnosť *Show spectrogram*. Ostatné možnosti umožňujú zobrazenia krivky základnej frekvencie (*Pitch*), hlasitosti (*Intensity*), priebehu formantov F1-F5 (*Formants*) a hlasivkových cyklov (*Pulses*).

Tento záznam reči znázorňuje niektoré akustické charakteristiky slova *konat*. Vidíme, že hláska /k/ sa začína úplne rovnou čiarou oscilogramu, teda nulovej amplitúdy, čo zodpovedá úplne bielemu poľu v spektrograme. Takýto začiatok majú všetky neznelé záverové spoluhlásky (v slovenčine /p/, /t/, /tʃ/, /k/ a tiež /c/, /č/), pretože cez ústnu dutinu neprúdi žiadny zvuk a výsledkom je ticho. Po uvoľnení uzáveru medzi zadnou časťou jazyka a horným mäkkým podnebí sa vzduch nahromadený za touto prekážkou náhle rozprúdi, čo zodpovedá krátkemu intervalu šumu s vysokou

frekvenciou a s aperiodickým priebehom. Priblížte si túto časť v PRAAT-e a presvedčte sa, že tento interval reči je naozaj aperiodickým šumom.

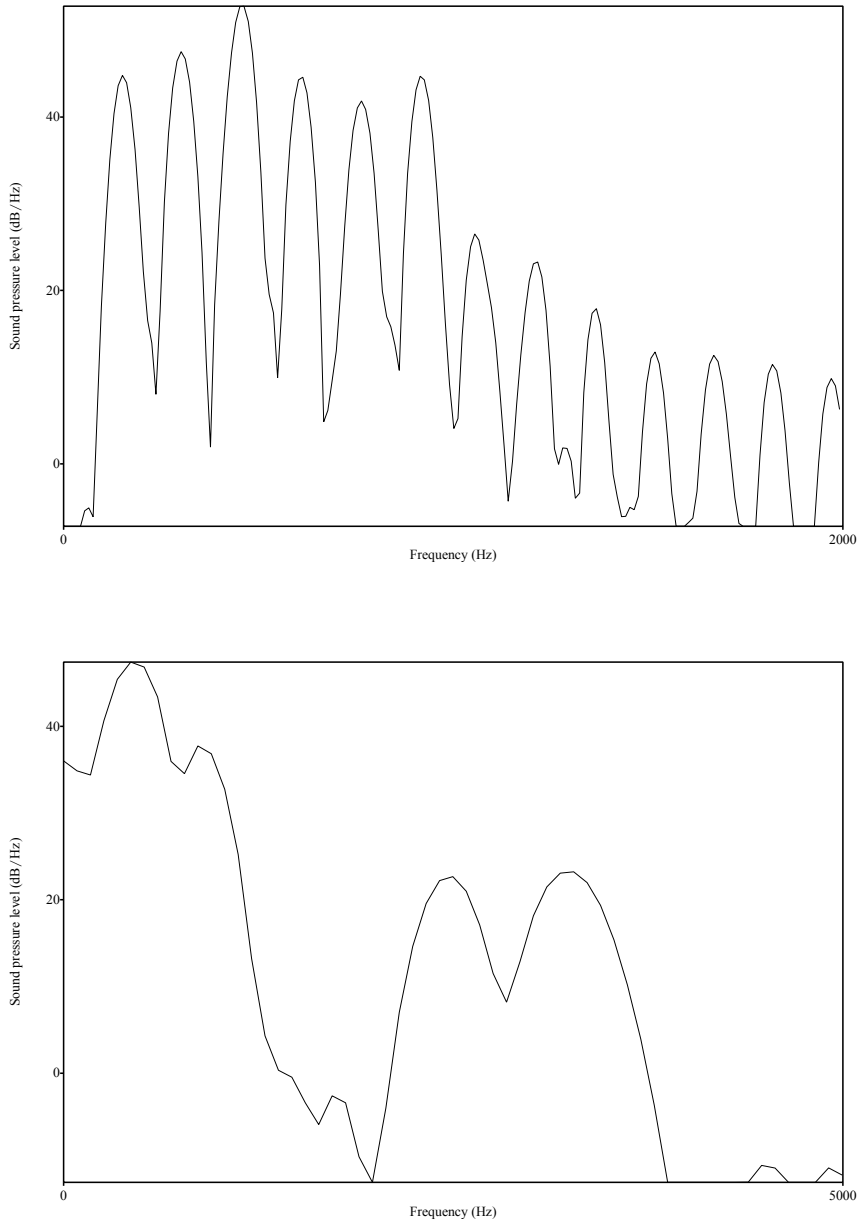
Ďalej nasleduje samohláska /o/. Na oscilograme vidíme rapídne zvýšenie amplitúdy a tým akustickej energie, a na spektrograme vidíme jej rozloženie do formantových frekvencií. Tmavé oblasti znázorňujú frekvencie s vysokou amplitúdou, a bledé oblasti frekvencie s nízkou. Približný priebeh prvých štyroch formantov je na Obrázku (9) naznačený bielymi bodkovanými čiarami.

Nasleduje spoluhláska /n/, ktorá má o niečo nižšiu akustickú energiu ako predchádzajúca samohláska, čo vidíme aj na osi y oscilogramu aj bledšími formantovými frekvenciami na spektrograme. Tento pokles energie je spôsobený záverom ústnej dutiny medzi špičkou jazyka a tvrdým podnebíom za zubami. Vzduch pri tejto spoluhláske uniká nosovou dutinou, čo je na spektrograme viditeľné typickou oblasťou „antirezonancie“, teda skoro bielej oblasti okolo frekvencie 1500Hz.

Spoluhláska /n/ potom prechádza do samohlásky /a/, ktorá má trochu viac akustickej energie ako /n/ ale oveľa menej v porovnaní s prvou samohláskou /o/. Je to preto, lebo /o/ je prízvuchné, keďže sa nachádza v prvej slabike slova, čo je slabika, ktorá je v spisovnej slovenčine nositeľom slovného prízvuku.

Slovo sa končí spoluhláskou /tʰ/, kde podobne ako pri /k/ vidíme krátky úsek nulovej energie počas úplného uzáveru medzi telom jazyka a tvrdým podnebíom, a potom aperiodickú vlnu po uvoľnení tohto uzáveru.

Na záver si ukážeme vzťah medzi spektrogramom a spektrom na príklade nášho slova *konatʰ*. Kliknite na *Sound konatʰ* v okne *Praat Objects* a kliknite *Edit*. Označte približne trvanie samohlásky /o/ (*Select* → *Select...*) od 0.1s po 0.17s. Teraz vytvorte spektrum tohto úseku (*Spectrum* → *View spectral slice*). Otvorí sa nové okno grafu frekvencia na osi x a intenzity na osi y. Označte myšou zhruba prvú štvrtinu okna a kliknite tlačidlo *Sel* vľavo dole, čím si priblížite úvodnú časť spektra pre frekvencie od 0 po zhruba 5000 Hz. a zistíte, že ukazuje viac-menej pravidelne sa opakujúce maximá (ak je potrebné priblížiť sa ešte viac, použite tlačidlo *In* vedľa tlačidla *Sel*). Nastavte kurzor na prvé maximum a na hornej lište sa objaví hodnota x okolo 150Hz, čo je základná frekvencia F0. Ďalšie maximum má hodnotu okolo 300Hz, teda dvakrát F0. Ďalšie maximum je okolo 450Hz, teda trikrát F0. Takže na tomto spektre vidíme F0, ktorá predstavuje frekvenciu kmitania hlasiviek hovoriaceho počas samohlásky /o/, spolu s jej alikvotnými frekvenciami. Horný panel Obrázku (10) zobrazuje tento graf pre interval frekvencií 0-2000Hz na osi x.



Obrázok (10) Skutočné spektrum pre samohlásku /o/ v slove *konat'* v hornom paneli a spektrum so zosilnenými formantovými frekvenciami v časovom bode 0.1283 sekundy v dolnom paneli.

Teraz zatvorte okno tohto spektra a v pôvodnom okne *Sound konat* na hornej lište kliknite *Formants* a vyberte prvú možnosť *Show formants*. Tmavé formantové frekvencie, hlavne pri samohláskach sa zvýraznia červenými bodkami. Nastavte základné parametre pre výpočet formantov pomocou *Formant* → *Formant setting...* . Môžeme ponechať všetky predvolené hodnoty, len je nutné ohraničiť hornú hodnotu pre výpočet formantov na 5000Hz u mužského hlasu a na 5500Hz u ženského. Keďže príklad slova *konat*, ktorý používame, bol nahratý mužským hlasom, uvedieme v prvom bode *Maximum formant (Hz)* hodnotu 5000. Teraz nastavte kurzor presne na čas 0.1283 (*Select* → *Move cursor to...*) a z možnosti *Formant* na lište vyberte *Formants listing*. V novom okne Praat ukáže hodnoty prvých štyroch formantov:

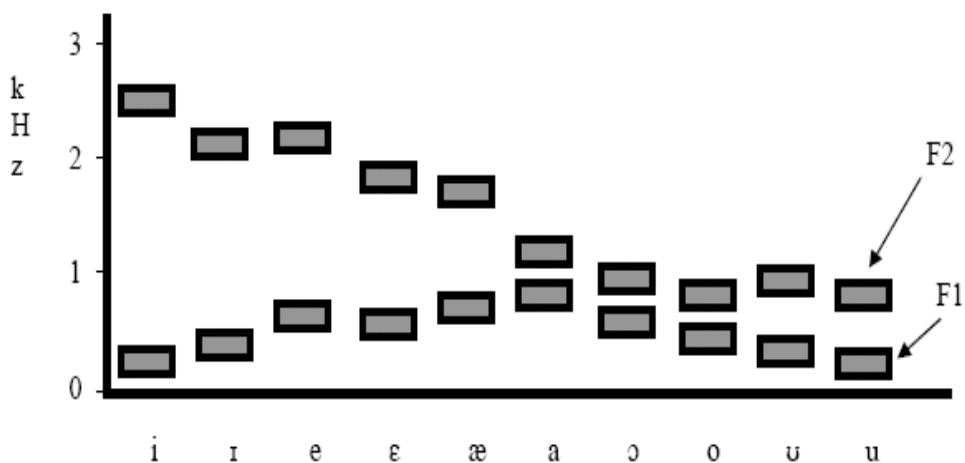
Time_s	F1_Hz	F2_Hz	F3_Hz	F4_Hz
0.128300	472.455153	897.821317	2505.970542	3259.198108

Ako posledný krok si vytvoríme spektrum z tohto časového bodu. Ak ste nepohli kurzorom, vyberte *Spectrum* → *View spectral slice*, a ak ste už kurzorom pohli, nastavte si ho predtým znovu na hodnotu 0.1283 (*Select* → *Move cursor to...*). V novom okne sa zobrazí spektrum z tohto bodu. Priložením kurzora na štyri hlavné maximá zistíme, že majú približne rovnakú hodnotu ako formanty uvedené vyššie. Dolný panel Obrázku (10) zobrazuje tento graf pre interval frekvencií 0-5000Hz na osi x. Obrázok (8) tak zobrazuje ilustračné a Obrázok (10) skutočné grafy akustického spektra ľudskej reči.

Štúdium formantových frekvencií je dôležité aj z toho dôvodu, lebo nám umožňuje relatívne verne odhadnúť polohu jazyka pri realizácii samohlások, ako je to naznačené na Obrázku (á). Prehľad referenčných hodnôt F1 a F2 pre hlavné anglické samohlásky je načrtnutý na Obrázku (11). Šedé obdĺžniky označujú zosilnené (tmavé) oblasti frekvenčného spektra ako je to na spektrograme na 0. Vidíme, že hláska [i] ako v slove *beet* má veľmi nízky F1 a veľmi vysoký F2 a artikulačne sa produkuje veľmi vysokou a prednou pozíciou jazyka. Hláska [a] má úplne odlišný formantový zápis: hodnota F1 je najvyššia spomedzi všetkých hlások, a hodnota F2 je zhruba v strede ostatných F2 hodnôt. Artikulačne je [a] najotvorenejšou hláskou, teda pozícia jazyka je najnižšie, a v horizontálnej rovine je trochu viac vzadu od stredu. Nakoniec, zadné vysoké hlásky /u/ majú nízky aj F1 aj F2. Všeobecne teda platí, že F1 frekvencia je negatívne priamo úmerná **vertikálnej** polohe jazyka: čím vyššia F1, tým nižšia poloha jazyka. Vidíme to na obrázku, keď smerom od [i] po [a] F1 stúpa a smerom od /a/ po /u/

opäť klesá. Druhý formant F2 je potom priamo úmerný **horizontálnej** polohe jazyka: čím vyššia F2, tým viac vpredu je poloha jazyka. Hlávka [i] je artikulovaná najviac vpredu a má najvyššiu F2, a tá postupne klesá až k minimálnym hodnotám pre hlávku [u], ktorá sa artikuluje najviac vzadu.

Pre nás je tento principiálny vzťah medzi merateľnými akustickými vlastnosťami vyjadrenými hodnotami formantov na jednej strane, a artikulačnými charakteristikami polohy jazyka a pier na strane druhej dôležitý z hľadiska relatívneho porovnávania výslovnosti na základe hypotéz interferencie materinského jazyka, ktoré popisujeme a analyzujeme v ďalších kapitolách. Napríklad, ak predpokladáme, že Slováci hovoriaci po anglicky nebudú rozlišovať medzi produkciou napätého [i] a nenapätého [ɪ], môžeme si vizuálne porovnať a kvantitatívne overiť, či F1 pre [i] je nižšie ako pre [ɪ], a či F2 pre [i] je vyššie ako pre [ɪ].



Obrázok (11) Referenčné hodnoty prvého (F1) a druhého (F2) formantu pre samohlávky anglického jazyka (prispôsobené z Ladefoged, 2001)

2.3. Záver

V tejto kapitole sme sa oboznámili so základmi akustiky reči a možnosťami softvéru PRAAT pri vizualizácii akustických vlastností reči. Začali sme jednoduchým fyzikálnym vyjadrením zvukovej vlny ako lokálneho znižovania a zvyšovania tlaku,

ktoré sa šíri plastickým prostredím ako napríklad vzduchom alebo vodou. Potom sme si charakterizovali fyzikálne vlastnosti týchto vln, a na záver uviedli príklady akustického záznamu ľudskej reči. Poslednú časť kapitoly tvorí stručný prehľad základných funkcií softvéru PRAAT pre vizualizáciu reči. Teraz sme v dobrej pozícii pristúpiť k popisu praktickej práce s nahrávkami reči, ktorý je hlavnou témou nasledujúcej kapitoly.