

# Primjena PRAAT programa u logopediji

---

Spajić, Kristina

Master's thesis / Diplomski rad

2019

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **University of Zagreb, Faculty of Education and Rehabilitation Sciences / Sveučilište u Zagrebu, Edukacijsko-rehabilitacijski fakultet**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:158:955530>

*Rights / Prava:* [In copyright](#) / [Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2024-09-11**



*Repository / Repozitorij:*

[Faculty of Education and Rehabilitation Sciences - Digital Repository](#)



Sveučilište u Zagrebu  
Edukacijsko–rehabilitacijski fakultet

Diplomski rad

**Primjena PRAAT programa u logopediji**

Kristina Spajić

Zagreb, lipanj 2019.

Sveučilište u Zagrebu  
Edukacijsko-rehabilitacijski fakultet

Diplomski rad

**Primjena PRAAT programa u logopediji**

Kristina Spajić

prof.dr.sc. Mladen Heđever

Zagreb, lipanj 2019.

### Izjava o autorstvu rada

Potvrđujem da sam osobno napisala rad (***Primjena PRAAT programa u logopediji***) i da sam njegova autorica.

Svi dijelovi rada, nalazi ili ideje koje su u radu citirane ili se temelje na drugim izvorima jasno su označeni kao takvi te su adekvatno navedeni u popisu literature.

Ime i prezime: Kristina Spajić

Mjesto i datum: Zagreb, lipanj 2019.

## ***Zahvale***

Zahvaljujem majci Jozefini, Onoj koja mi je omogućila školovanje i jedinoj koja mi je htjela dobro više nego sebi. Hvala Ti što si mi pružila sve što si mogla i više nego što si mogla. Nisi doživjela u ovoj stvarnosti trenutak moje diplome, ali znam da si prisutna na neki drugi način, onako kako samo majke znaju. Iznenada mi te fizička stvarnost ukrala, ali te je ostavila ona druga vrsta stvarnosti i prisutnosti.

Legendo, volim Te i hvala Ti na svemu.

Zahvaljujem i svojoj drugoj majci, svojoj Babi iz Konavala, koja je neprestano provjeravala jesam li više počela pisati diplomski, da završim više ovu fakultetsku priču nakon svih nesreća i koja mi je isto htjela sve najljepše i najbolje što se poželjeti može. Ni Ti nisi doživjela završetak ovog diplomskog rada, ali znam da svejedno sve vidiš i čuješ.

Ovo je i za Tebe.

Hvala i mentoru na strpljenju i razumijevanju.

## Primjena PRAAT programa u logopediji

Kristina Spajić

Prof.dr.sc. Mladen Heđever

Sveučilište u Zagrebu, Edukacijsko-rehabilitacijski fakultet, Odsjek za logopediju

### **Sažetak**

Logopedija je interdisciplinarna znanost na raskrižju čiste i primijenjene znanosti. Odlikuje je brojnost i raznolikost područja kojih se dotiče, a jedno od užih područja je i područje akustike. U domeni logopedije, akustika kao znanost najveću ulogu ima u području poremećaja glasa. Na tržištu postoje brojni programi za analizu glasa i govora koji su dostupni i priznati od akademske zajednice. Jedan od njih je i program PRAAT. To je besplatni računalni softverski paket namijenjen znanstvenoj analizi govora. Koristi se u fonetici, akustici, forenzičkoj akustici i fonetici te logopediji. Neprestano se revidira i poboljšava novim verzijama, a dostupan je većini operativnih sustava. Program nudi široki raspon funkcija i analiza, uključujući spektrogramsku analizu, analizu parametara glasa, analizu formanta itd. Jedna od mogućnosti koje program nudi je i funkcija PRAAT Scripts, odnosno programski jezik koji omogućava generiranje i automatiziranje sekvenci koje bi se inače morale raditi korak po korak. U radu je u cijelosti prikazan AVQI Script te njegov značaj u dijagnostičke svrhe. Najveća prednost PRAAT-a je u tome što uvijek pruža slobodan pristup najnovijim verzijama samog softvera. Solidno poznavanje ovog programa daje dugotrajne dobrobiti kliničarima, znanstvenicima, a naposljetku i studentima logopedije koji pokazuju veće zanimanje u područjima akustike i fonetike.

***Ključne riječi:*** akustika, PRAAT, Script, logopedija

## Application of PRAAT Programme in Speech-Language Pathology

Kristina Spajić

Prof.dr.sc. Mladen Heđever

University of Zagreb, Faculty of Education and Rehabilitation Sciences, Department of  
Speech and Language Pathology

### Summary

Speech-Language Pathology is an interdisciplinary science at the intersection of pure and applied science. It is distinguished by the number and the variety of areas it touches, and one of the narrower areas is the area of acoustics. In the domain of Speech-Language Pathology, acoustics as a science plays a major role in the field of voice disorders. There are numerous voice and speech analysis programmes available and recognized by the academic community on the market. One of them is the PRAAT programme. It is a free software package designed for scientific analysis of speech. It is used in phonetics, acoustics, forensic acoustics and phonetics and speech-language pathology, as well. It is constantly revised and improved with newer versions, and is available for most operating systems. The programme offers a wide range of functions and analyses, including spectrogram analysis, voice parameters analysis, formant analysis etc. One of the features that the programme offers is also the PRAAT Scripts function, which is the programming language that allows generating and automation of sequences that would otherwise have been done step by step. In this paper the AVQI Script is fully depicted and its significance for diagnostic purposes, as well. PRAAT's greatest advantage is that it always provides free access to the latest versions of the software itself. Solid knowledge of this programme provides long-term benefits to clinicians, scientists, and finally to speech-language pathology students who show greater interest in the fields of acoustics and phonetics.

**Key words:** acoustics, PRAAT, Script, Speech-Language Pathology

## SADRŽAJ:

|                                           |    |
|-------------------------------------------|----|
| 1. UVOD.....                              | 1  |
| 1.1.    Određenje akustike .....          | 2  |
| 1.2.    Govorna akustika.....             | 2  |
| 1.3.    Procesiranje govora .....         | 4  |
| 1.3.1.    Kodiranje govora.....           | 4  |
| 1.3.2.    Nyquistov teorem .....          | 7  |
| 1.4.    Odrednice zvuka .....             | 8  |
| 1.4.1.    Frekvencija .....               | 8  |
| 1.4.2.    Valna duljina.....              | 8  |
| 1.4.3.    Intenzitet.....                 | 9  |
| 1.4.4.    Boja (timbar) .....             | 9  |
| 1.5.    Odrednice glasa.....              | 10 |
| 1.5.1.    Fundamentalna frekvencija ..... | 10 |
| 1.5.2.    Harmonici.....                  | 10 |
| 1.5.3.    Formanti.....                   | 12 |
| 1.5.4.    Jitter.....                     | 12 |
| 1.5.5.    Shimmer.....                    | 12 |
| 1.6.    Spektar govora .....              | 13 |
| 1.6.1.    FFT spektar .....               | 13 |
| 1.6.2.    LPC spektar .....               | 14 |
| 1.6.3.    Spektogram .....                | 16 |
| 2.    CILJ RADA .....                     | 17 |
| 3.    PRAAT .....                         | 18 |
| 3.1.    Osnovni principi rada.....        | 20 |
| 3.1.1.    Pokretanje programa .....       | 20 |
| 3.1.2.    Analiza zvučnog zapisa .....    | 22 |
| 3.2.    PRAAT Script.....                 | 25 |
| 3.2.1.    AVQI SCRIPT.....                | 26 |
| 4.    PRAAT U LOGOPEDIJI.....             | 30 |
| 5.    ZAKLJUČAK.....                      | 32 |

|                            |    |
|----------------------------|----|
| 6. LITERATURA .....        | 33 |
| POPIS SLIKA I TABLICA..... | 37 |

## 1. UVOD

Logopedija je interdisciplinarna znanost na raskrižju čiste i primijenjene znanosti. U sebi objedinjuje znanja iz medicine, lingvistike i fonetike, psihologije i odgojnih znanosti te pruža procjenu, tretman, podršku, prevenciju i savjetovanje za djecu i odrasle koji imaju bilo koji oblik teškoća u komunikaciji, govoru te hranjenju i gutanju. Najpoznatiju i sveobuhvatnu definiciju logopedije daje nam Europsko udruženje logopeda – Comité Permanent de Liaison des Orthophonistes / Logopedes de l'Union Européenne (CPLOL), a ona glasi: *Logopedija je znanost koja se bavi prevencijom, otkrivanjem, dijagnosticiranjem i tretmanom poremećaja humane komunikacije pod kojom se podrazumijevaju svi oni procesi i funkcije koji su povezani s produkcijom govora te s percepcijom i produkcijom oralnog i pisanog jezika, kao i oblicima neverbalne komunikacije.* Iz navedenog može se vidjeti kako je logopedija kao znanost uistinu raznolika. Na sam spomen pojma logopedije velik broj osoba reći će kako ih to najprije podsjeća na mucanje, ispravljanje pogrešnog izgovora (najčešće glasa „R”) te disleksiju. Akustika glasa i govora jedno je od užih područja kojima se logopedija bavi, a ujedno će određene stavke iz područja akustike biti obuhvaćene ovim diplomskim radom s ciljem osvještavanja kako logopedija nije samo „R”. Rad će obraditi neke pojmove iz akustike u kontekstu računalnog programa PRAAT, kako se on koristi za osnovne funkcije koje mogu biti informativne i korisne i nekome tko nije inicijalno logopedске struke. Nadalje, opisat će se kako koristiti ovaj program u logopediji te važnost njegove uloge. Cilj je da ovaj rad bude svojevrsan „brzinski” priručnik za bilo koga tko želi znati malo više o ovom programu.

## **1.1. Određenje akustike**

Govor se kao zvučna pojava počinje proučavati još prije II. Svjetskog rata iako povijest akustike seže prilično ranije. Postoje povijesni podatci da su se prva akustička mjerenja pojavila već u 16. i 17. stoljeću. Fiziološka i govorna akustika kao grane akustike proučavaju kako proizvodimo (govorni) zvuk i kako ga nadalje percipiramo. Osnovni zvuk glasa nastaje u larinksu prolaskom zračne struje i istovremenim treperenjem glasnica (Heđever, 2012). Djelovanjem četiri procesa ostvaruje se govor kao zvučna pojava, a to su: respiracija, fonacija, rezonancija i artikulacija. Njihovim istovremenim i ujednačenim djelovanjem nastaje govor (Škarić, 1991). Govor kao zvučna pojava predmet je zanimanja brojnim stručnjacima. Sve više se javlja potreba za raznovrsnom digitalnom obradom govora i koliko su brojna područja u kojima se primjenjuju postupci digitalne obrade govora. Neka od tih područja su sljedeća: fonetika, forenzička akustika i fonetika, telekomunikacije, multimedijски sustavi, digitalna televizija i radio, informacijske i obavještajne službe te naposljetku i logopedija. Zbog širokog dijapazona ukupne problematike, poznavati i koristiti svaku moguću primjenu unutar svakog navedenog područja je praktički nemoguće. Stoga, stručnjaci se profiliraju u različitim područjima upotrebe ovisno o vlastitim fokusima ili vlastitog interesa ili vlastite struke (Petrinović, 2002).

## **1.2. Govorna akustika**

Tiwari (2012) iznosi da humana vokalizacija predstavlja zvukove koje može proizvesti iznimno ljudski govorni trakt. Među ostalim vokalizacijama, odnosno glasanjima poput kašljanja, kihanja, smijanja i vrištanja, govor je najvažniji. To je sasvim razumljivo zato što je govor primarni medij sofisticiranog sustava ljudske simboličke komunikacije – jezika. Govor je zapravo učinio jezik zvučno ostvarivim, odnosno sve ono što je u govornikovim mislima u obliku jezične forme.

Za nastanak glasa neophodna su tri elementa: zračna struja, glasnice i rezonatori. Zračna struja nastaje u plućima. Glasnice su generator glasa čijim treperenjem nastaje čujni

zvuk, a rezonatori su šupljine koje osnovnom laringealnom tonu daju čujnost, boju i glasnoću. Nadalje, izgovorni organi ili artikulatori su oni koji moduliraju, odnosno oblikuju zračnu struju čime se postiže realizacija ostalih glasova u govoru. Kao što je već prethodno bilo spomenuto, govor je pojava koja se ostvaruje koordiniranim radom respiracije, fonacije, rezonancije i artikulacije. Govorni trakt se može stoga promatrati kao akustički sustav koji definiraju njegove anatomske i funkcionalne osobine, a na osnovu tih osobina moguće je prikazati akustičke principe generiranja govora. Kako bi se govor uopće i ostvario, ekspiriranu zračnu struju potrebno je modulirati s dovoljno brzim promjenama kako bi one postale čujne i naposljetku percipirane (Heđever, 2012). Prema Heđever (2012; str. 96) „*Kod govora razlikujemo četiri vrste modulacija:*

- *modulaciju glasnicama*
- *modulaciju trenjem (frikcijska modulacija)*
- *start-stop modulaciju (okluzija)*
- *modulaciju rezonantnim šupljinama.*”

Govor je, dakle, složeni sustav raznolikih segmenata koji se odvijaju simultano u vremenu. Međutim, elemente govora treba promatrati kao segmente koji su dio veće cjeline, a analizom tih pojedinih segmenata dobivamo detaljnije iskoristive informacije izoliranih segmenata uzetih iz kontinuuma (Heđever, 2012).

Svaki govorni iskaz ima svoju formu i sadržaj. Sadržaj se odnosi na sve ono što je jezično u svojoj biti i da se zapravo radi o poremećaju jezika. Kao dobar primjer poslužiti će Wernickeova ili receptivna afazija koja je dobila ime po njemačkom neuropsihijatru Carlu Wernickeu (1848 – 1905) koji je prvi opisao ovaj tip afazije. On navodi kako su pacijenti govorili tečno i gramatično, ali sadržaj njihovih iskaza nije imao nikakvoga smisla. Također, imali su priličnih teškoća u razumijevanju govorenog jezika (Hegde, 2006). Ukoliko je poremećena forma govornog iskaza, sasvim sigurno se radi o tome da je poremećen neki od četiri prethodno spomenuta procesa koji čine formu govora – respiracija, fonacija, rezonancija i artikulacija. Akustičkom analizom govora moguće je utvrditi koji su procesi sudjelovali u realizaciji govora, ali moguće je odrediti i za svaki pojedini segment govornoga zvuka (Heđever, 2012).

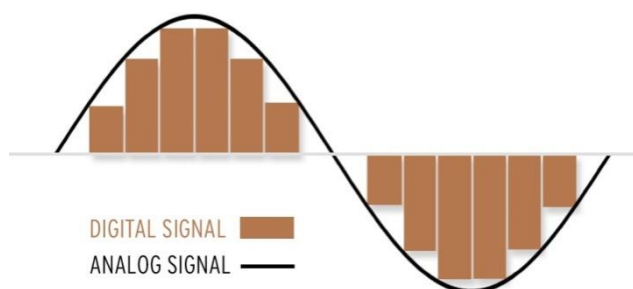
### 1.3. Procesiranje govora

Tehnološki napredak i novije tehnologije dovele su do značajnog razvoja u području govornog procesiranja. „Govorno procesiranje (*speech processing*, obrada govora) proučava sljedeće aspekte:

1. Prepoznavanje govora (*speech recognition*), bavi se analizom jezičnog sadržaja govornog signala,
2. Prepoznavanje govornika (*speaker recognition*), bavi se prepoznavanjem govornika (pojedine osobe) na temelju njegova govora,
3. Kodiranje govora (*speech coding*), bavi se obicima kompresije (sažimanja) govornog signala, važno je u telekomunikacijskom području,
4. Sintetiziranje govora (*speech synthesis*): stvaranje umjetnog govora, najčešće pomoću računala,
5. Poboljšanje razumljivosti govora (inteligibilitet), odnosi se na unapređenje razumljivosti i perceptivne kvalitete govornog signala (npr. smanjenje šuma audio signala, filtriranje audio signala i sl.)" (Hedevar, 2012, str. 14).

#### 1.3.1. Kodiranje govora

Kodiranje govora intenzivno se razvija i koristi u posljednjem kvartalu 20. stoljeća. Tom razvoju najviše je pridonijelo uvođenje digitalne tehnologije u istraživanja. Sam pojam digitalizacije odnosi se na pretvaranje analognog u digitalno, odnosno u ovom slučaju pretvaranja akustičkog govornog signala iz analognog u digitalni oblik, a tri parametra koja su neizostavna u tom procesu su: uzorkovanje (*sampling*), kvantiziranje (*quantization*) i kodiranje (*coding*).



**Slika 1:** Analogni i digitalni signal (Preuzeto iz Gans, 2015).

Vizualnu razliku između analognog i digitalnog signala prikazuje Slika 1. Gans (2015) ovu razliku opisuje na način da se digitalni val mora „penjati i spuštati niz stepenice“ , što je suprotno od „ležernog kotrljanja preko brijegova“ analognog signala. Ova usporedba ide u prilog analognome signalu, međutim, kada se uzmu u obzir sve prednosti i nedostaci oba signala, današnja tehnologija ipak više podržava pohranu informacija u digitalnom obliku.

Analogno-digitalna pretvorba (digitalizacija) odnosi se na pretvaranje informacije sadržane u nekoj električnoj veličini iz analognog u digitalni (brojčani) oblik. Sam pojam analogan znači da je to nešto što je svojstveno sustavu gdje se informacije ili podatci prikazuju fizikalnom veličinom koja unutar određenih granica može poprimiti bilo koju vrijednost. Suprotno tomu je pojam digitalan. Pojam digitalan svojstven je sustavu u kojem se podatci iskazuju fizikalnom veličinom koja unutar točno određenih granica može poprimiti samo neku od unaprijed određenih vrijednosti, npr. u binarnom sustavu 0 ili 1. Važno je naglasiti da su te vrijednosti diskretne (Leksikografski zavod Miroslav Krleža <http://www.enciklopedija.hr/natuknica.aspx?ID=2475>).

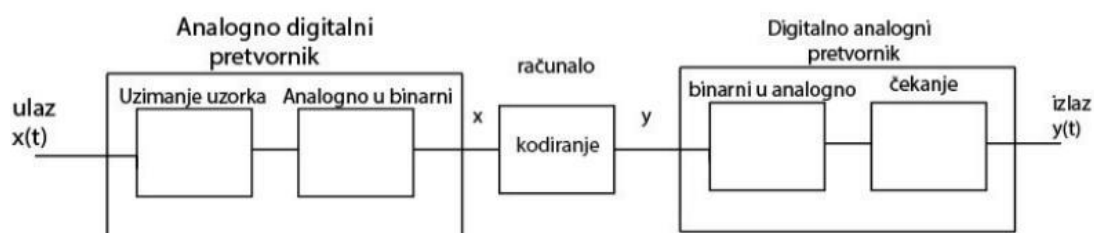
Prednost digitalne informacije nad analognom je u tome što je digitalna informacija pri obradi manje podložna smetnjama, oštećenjima, izobličavanjima i inim utjecajima. Pretvorba se vrši uzorkovanjem analogne veličine u vremenskim odsječcima te kvantiziranjem i kodiranjem tih uzoraka (Leksikografski zavod Miroslav Krleža <http://www.enciklopedija.hr/natuknica.aspx?ID=2475> ).

Uređaji za digitalizaciju mjere vrijednosti električnoga signala u određenim, kratkim vremenskim intervalima (uzorkovanje); izmjerene vrijednosti zaokružuju (kvantizacija) te ih zapisuju u obliku binarnoga koda. U kontekstu akustičkih informacija, odnosno zapisa, kvaliteta digitalnog zapisa ovisi o količini bitova kojima se opisuje zvuk tijekom jediničnoga vremenskog odsječka (engl. *bit rate*), obično sekunde (bit/s ili bps - *bits per second*; bitova po sekundi), o čemu ponajviše i ovisi sama kvaliteta digitalnoga zapisa. To otvara mogućnost nastanka različitih formata zvučnih zapisa, a na računalima su najčešći WAV, WMA i MP3 format (Frosch, 2017). Najvažniji preduvjet kvalitete jest da frekvencija uzorkovanja mora biti dvostruko veća od najviše frekvencije analognog signala kako bi se izbjegli gubitci (Heđever, 2012). Digitalni zvučni zapis determiniran je trima parametrima: brojem kanala (mono ili stereo), frekvencijom uzorkovanja te rezolucijom (odnosi se na obuhvaćanje razina glasnoće). Optimalna frekvencija uzorkovanja kod zvučnih zapisa iznosi 44100 Hz, odnosno

ona je uobičajena vrijednost za kvalitetno snimanje zvuka (Heđever, 2012). Ovdje važnu ulogu ima koncept Nyquistova teorema, o kojemu će više biti riječ u sljedećem poglavlju.

Smisao kodiranja jest pohraniti i reproducirati zapis sa što je moguće manjim brojem bitova, a pritom se uglavnom temelji na svojstvima kvantiziranog signala i osnovama psihoakustike. Ulazni signal razlaže se na frekvencijske komponente i prenosi informaciju o svakoj pojedinoj komponenti, tj. frekvenciju, amplitudu, fazu i trajanje. Ovaj postupak je vrlo uspješan kod prijenosa čistih tonova, odnosno zvučnih signala koji se većinom sastoje od harmonički ovisnih komponenti ([https://www.fer.unizg.hr/download/repository/PrZv\\_10\\_2018.pdf](https://www.fer.unizg.hr/download/repository/PrZv_10_2018.pdf)).

Računalo pohranjuje zapis u digitalnom obliku s konačnim brojem bitova. Svaki uzorak analognog signala naposljetku treba pretvoriti u konačan broj bitova prije nego što može ići na daljnju obradu upravo pomoću računala. Ta se pretvorba postiže analogno-digitalnim pretvornikom (ADC – *Analog – digital converter*); pogledati Sliku 2. Ključan parametar ovog pretvornika jest njegova rezolucija, odnosno upravo prethodno spomenuta količina bitova koji koristi za prikazivanje analognog signala. S obzirom da je broj bitova konačan, a kontinuirani analogni signal može imati beskonačno mnogo vrijednosti, tada je potrebno određeni niz analognog signala mapirati na određenu kombinaciju bitova. Zaključno – što je veća rezolucija analogno-digitalnog pretvornika, to je i veći broj dijelova na koje je moguće mapirati pojedine vrijednosti signala, a samim time detektirati manje promjene mjernog signala (Pinculić, 2015).



**Slika 2.** Općeniti princip digitalne obrade signala (Preuzeto iz: Pinculić, 2015).

### 1.3.2. Nyquistov teorem

Osnovna karakteristika procesa kvantizacije jest pojava greške, odnosno prisustva šuma u signalu u kontekstu akustike. Utjecaj šuma kvantizacije ovisi o rezoluciji, to jest o veličini koraka kvantizacije. Međutim, postoji nešto što omogućuje smanjenje broja šumova odnosno gubitka kvalitete u zapisu. Teorem uzorkovanja ili Nyquistov teorem (punog naziva Nyquist-Shannonov teorem) pokazuje kako se kontinuirani analogni signal može pretvoriti u slijed vremenski diskretnih impulsa i obrnuto, tako da nema nikakvog gubitka informacija, ukoliko je ostvareno sljedeće: frekvencija uzorkovanja ( $f_{sr}$  - *sampling rate*) kojom se uzimaju uzorci mora biti najmanje dva puta veća (viša) od najviše frekvencije  $f_m$  koja se može pojaviti u analognom signalu koji se uzorkuje. Ovaj teorem dobio je ime po američkom fizičaru i inženjeru elektrotehnike i računarstva švedskog podrijetla Harryju Nyquistu (1889 – 1976) koji je svoj najveći doprinos dao u području telekomunikacija (Encyclopaedia Britannica).

Matematičkim jezikom, to prikazuje ovaj izraz  $f_{sr} \geq 2f_m$  (Pinculić, 2015). Dakle, u analognom kontinuiranom signalu ne smije biti viših frekvencija nego što je polovica frekvencije uzorkovanja, kako bi sama pretvorba iz kontinuiranog signala u niz diskretnih impulsa bila ostvariva bez ikakvog gubitka informacija. To se odnosi i na obrnuti postupak, pretvorbu iz digitalnog signala u analogni (Pinculić, 2015). Odnos između frekvencije uzorkovanja i kvalitete zvuka vjerno prikazuje Tablica 1.

**Tablica 1.** Frekvencija uzorkovanja i kvaliteta zvuka (Preuzeto iz Heđever, 2012).

| frekvencija uzorkovanja<br>(sempliranja) | gornja frekvencija<br>audio signala | napomena                |
|------------------------------------------|-------------------------------------|-------------------------|
| 8.000 Hz                                 | 4.000 Hz                            | telefonski prijenos     |
| 32.000 Hz                                | 16.000 Hz                           | cijelo govorno područje |
| 44.100 Hz                                | 22.050 Hz                           | cijelo čujno područje   |
| 96.000 Hz                                | 45.500 Hz                           | profesionalna uporaba   |

## **1.4. Odrednice zvuka**

Zvuk je pojava koju percipiramo osjetilom sluha. Fizikalne (objektivne) odrednice zvuka su frekvencija, intenzitet i trajanje. Zvuk je definiran i oblikom zvučnog vala, a te objektivne parametre doživljavamo subjektivno sluhom kao glasnoću, visinu, trajanje i boju (Heðever, 2012).

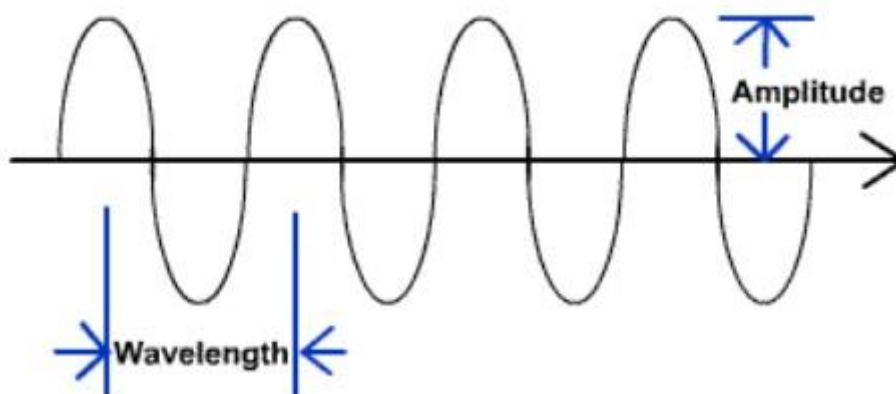
### **1.4.1. Frekvencija**

Frekvencija je fizikalna veličina kojom se izražava koliko se puta ponovio neki periodički događaj u određenom vremenskom intervalu. Mjerna jedinica je Hertz (Hz). Naziv je dobila po njemačkom fizičaru Heinrichu Rudolphu Hertzu (1857 – 1894). Ljudsko uho čuje zvukove u frekvencijskom rasponu od 16 do 20 000 Hz. Općenito, zvukove možemo podijeliti u šumove i tonove. Šum nastaje nepravilnim titranjem zvučnog izvora pri čemu se neprestano mijenjaju i frekvencija i amplituda te ispunjava cijelo jedno frekvencijsko područje (Lukavečki, 2008). Ton nastaje pravilnim, periodičnim titranjem zvučnog izvora i frekvencija je stalna. Ton je zapravo jednostavno sinusoidno titranje (Stone, 2001).

Visina tona koja je subjektivno doživljena ovisna je o frekvenciji zvučnoga vala. Zvukove visoke frekvencije doživljavamo kao „svjetlije“ i više, a zvukove niskih frekvencija doživljavamo kao „tamnije“ i dublje (Heðever, 2012).

### **1.4.2. Valna duljina**

Valna duljina periodičnoga vala jest udaljenost nakon koje se oblik vala ponavlja. Označava se grčkim slovom lambda ( $\lambda$ ). Jedan puni val predstavlja i jedan puni titraj čestice, a iz toga slijedi da onoliko valova koliko ima u jednoj sekundi zapravo određuje i frekvenciju tog vala (Lukavečki, 2008).



**Slika 3.** Valna duljina periodičnoga vala

(Preuzeto s [http://www.izor.hr/c/document\\_library/get\\_file?uuid=7eddcdd0-58af-4aba-bbe6-232a744be89b&groupId=31930](http://www.izor.hr/c/document_library/get_file?uuid=7eddcdd0-58af-4aba-bbe6-232a744be89b&groupId=31930)).

### 1.4.3. Intenzitet

Energija koju val prenese kroz jedinicu površine u jedinici vremena jest intenzitet vala nekog zvuka. Intenzitet zvuka subjektivno doživljavamo u logaritamskoj skali (Fechnerov zakon). Zato se uveo pojam razine intenziteta zvuka, koji je definiran kao deseterostruki logaritam omjera intenziteta zvuka i intenziteta zvuka na pragu čujnosti (referentnog zvučnog tlaka od  $20 \mu\text{Pa}$  (Heđever, 2012).

### 1.4.4. Boja (timbar)

Timbar (franc. *le timbre*) je svojstvo tona po kojem je uho sposobno razlikovati tonove iste visine, jačine i boje. Timbar ponajviše ovisi o svojstvu tvari koja rezonira, o stupnju jačine osnovnog tona, amplitudi, strukturi i sklopu jačine glavnih alikvotnih tonova određenog izvora zvuka, kao i o frekvenciji tona. Najširi dijapazon timbra ima ljudski glas (<http://proleksis.lzmk.hr/57191/>).

Najbolji primjer za razumijevanje koncepta timbra je kada bismo, primjerice, na dva potpuno različita instrumenta ili dva srodna instrumenta odsvirali određeni ton, npr. Fis (F $\sharp$ ). Flauta će izvesti zadani ton na sebi svojstven način, violina na svoj, a violončelo na svoj. Svi

ti instrumenti na istoj frekvenciji daju različite valne oblike, stoga ih raspoznavamo kao različite i sebi svojstvene.

Ako žica koja je određena frekvencijom od 220 Hz vibrira, generirani zvuk također sadrži komponente frekvencija 440 Hz, 660 Hz, 880 Hz, 1100 Hz i tako dalje. Iako slušatelji pretežno percipiraju osnovnu notu, intenziteti tzv. nadtonova definiraju zapravo karakter glazbenog instrumenta (Beer, 1998).

## **1.5. Odrednice glasa**

### **1.5.1. Fundamentalna frekvencija**

Glas je složeni zvuk. To znači da se sastoji od osnovnog tona i nadtonova ili harmonika. Njegovo titranje je periodično, tj. u svakom je dijelu spektra jednako. Ako bi se i uklonio osnovni ton, i dalje bi se čula ista visina glasa (Heđever, 2012). Fundamentalna frekvencija ( $F_0$ ), *pitch*, nulti harmonik ili osnovni laringealni ton je zvuk koji nastaje titranjem glasnica u larinksu, odnosno ona je osnovni segment glasa i označava prosječnu brzinu titranja glasnica tijekom govora. Kod muškaraca ona iznosi tijekom govora 100-150 Hz, kod žena 180 – 220 Hz, a kod djece iznosi oko 300 Hz (Heđever, 2012). Osnovni ton ne čuje se u govoru, ali zato njegovi viši harmonici dobivaju konačnu glasnoću zahvaljujući rezonanciji vokalnog trakta. Fundamentalna frekvencija laringealnoga tona ovisi o veličini i širini larinksa, duljini glasnica te djelomično o debljini glasnica (Titze i Talkin, 1979; prema Heđever, 2012).

### **1.5.2. Harmonici**

U svakom složenom harmoničnom zvuku kao cjelobrojni umnošci osnovnog tona (fundamentalne frekvencije) javljaju se harmonici i izrazito su vidljivi u svim vokalima, kao i u spektru zvuka koji proizvede neki glazbeni instrument. U ljudskom je glasu vidljivo i do 15 harmonika, a njihova frekvencija je određena fundamentalnom frekvencijom (primjerice, ako fundamentalna frekvencija iznosi 100 Hz, iz toga slijedi da prvi harmonik ( $H_1$ ) iznosi 200 Hz,

drugi harmonik (H2) 300 Hz, treći harmonik (H3) 400 Hz itd. (Heđever, 2012). Služeći se teorijom glazbe, glazbeni naziv za harmonike jesu alikvotni tonovi. Alikvotni tonovi predstavljaju harmonične komponente generiranog tona. Različiti brojevi postojećih alikvota i njihov različit relativni intenzitet unutar cjelokupnog zvuka određuju timbar zvuka, koji je prethodno opisan. Alikvotni se također nazivaju i parcijalnim, harmonijskim ili gornjim tonovima, odnosno parcijalama (Milivojević i sur., 2014).

Raspodjelom jačine tih alikvota zapravo ovisi boja nekog instrumenta, što matematički savršeno objašnjava *Prvi Pitagorin zakon malih brojeva* koji glasi: *Dva tona su konsonantna ako im frekvencije stoje u odnosu malih prirodnih brojeva*. Znači, tonovi koji su u harmoniji sa osnovnim tonom su upravo tonovi čije su frekvencije zastupljene u spektru alikvotnih tonova (Culek, 2015).

Svaki alikvot ili harmonik jest sinusoidalni ton čija je frekvencija točno određena ovisno o tome kolika je fundamentalna frekvencija tona. Svaki složeni ton sastavljen je od niza sinusoidalnih tonova koji tvore alikvotni niz koji se sastoji od fundamentalne frekvencije i pripadajućih joj viših harmonika. Taj determinirani odnos između fundamentalne frekvencije i svakog sljedećeg harmonika u glazbenom smislu prikazuje Tablica 2. na primjeru tona a, odnosno počevši od njega. Za svaki sljedeći član frekvencija je cjelobrojni viši iznos – 1:2:3:4:5:6:7.... itd. Dakle, boju tona određuju njegovi viši harmonici, a visinu tona određuje osnovni harmonik, odnosno fundamentalna frekvencija. Vokalni trakt (koji tvore sve šupljine iznad larinksa) zbog efekta rezonancije daje konačnu boju i intenzitet zvuka. Modulacija glasnicama stvara zvuk koji ima obilježja linijskoga frekvencijskog spektra, odnosno harmoničnoga zvuka (Culek, 2015).

Proučavajući akustičke fenomene, harmonike je prvi uočio francuski matematičar i glazbeni teoretičar Marin Mersenne (1588 – 1648). Prvi je definirao ton i zvuk kao titranje zraka te otkrio postojanje alikvotnih tonova (Lukavečki, 2008).

**Tablica 2.** Alikvotni tonovi iznad tona a (Preuzeto s <http://plavi.forumsr.com/t128p10-sve-o-klaviru>).

| Osnovni ton   | Alikvotni tonovi (Hz) |                |                |                  |                |                |                |                |                  |                  |                |                |
|---------------|-----------------------|----------------|----------------|------------------|----------------|----------------|----------------|----------------|------------------|------------------|----------------|----------------|
| <b>110 Hz</b> | 220                   | 330            | 440            | 550              | 660            | 770            | 880            | 990            | 1.100            | 1.210            | 1.320          | 1.430          |
| Približan ton | a                     | e <sup>1</sup> | a <sup>1</sup> | cis <sup>2</sup> | e <sup>2</sup> | g <sup>2</sup> | a <sup>2</sup> | h <sup>2</sup> | cis <sup>3</sup> | dis <sup>3</sup> | e <sup>3</sup> | f <sup>3</sup> |

### 1.5.3. Formanti

Formanti su intenzitetski naglašeni dijelovi spektra koji su rezultat rezonancije u rezonantnim šupljinama i ne ovise o promjenama visine osnovnog tona ili harmonika. Kod zvučnih glasova formante ne možemo uočiti direktno u spektru zbog prisutnosti harmonika, stoga se formanti uglavnom promatraju u izgovoru samoglasnika. Formanti ujedno predstavljaju pojačane intenzitetske vrijednosti grupiranih harmonika (Heđever, 2012). Kako bi se prikazali formanti glasa, služi LPC analiza koja će kasnije biti opisana.

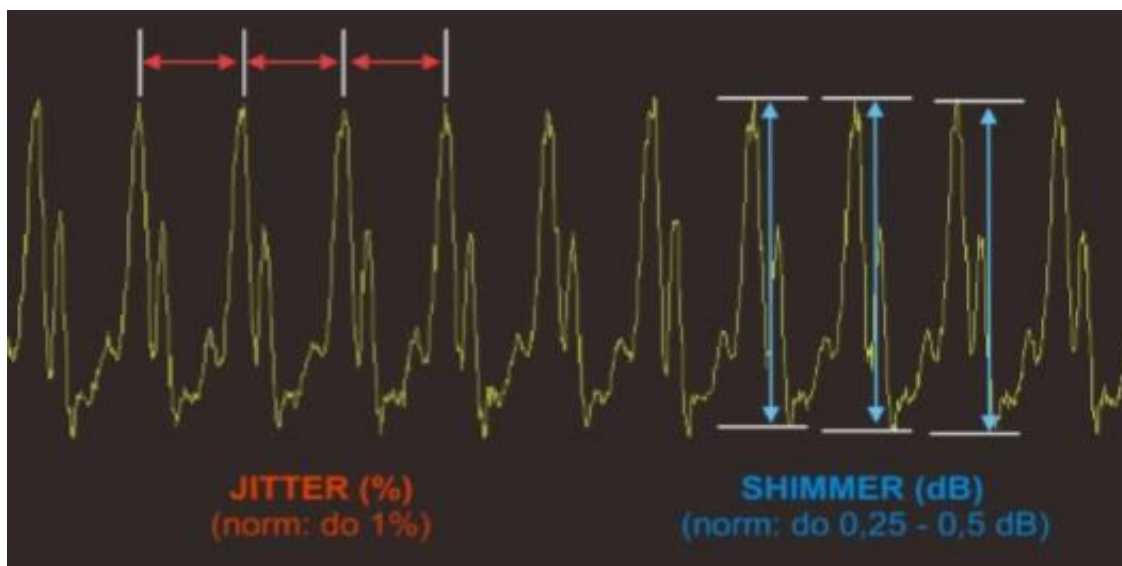
Analiza formanta na vokalima jedna je od najvažnijih metoda u glasovnim usporedbama, pogotovo u forenzičke svrhe (Lindh, 2017).

### 1.5.4. Jitter

*Jitter* je termin koji se odnosi na nepravilnosti u brzini vibracija glasnica, odnosno varijacije frekvencije osnovnog laringealnog tona. Najčešće se izražava kao prosječno odstupanje u brzini titranja glasnica u postotcima. Normalne vrijednosti *jittera* dopuštaju odstupanje do 1% (npr. ako je  $F_0 = 100$ , *jitter* će iznositi oko 1 Hz). Povišeni *jitter* manifestirat će se kao glas lošije kvalitete, a manja vrijednost *jittera* daje kvalitetniji glas kojemu je intonacija čišća (Heđever, 2012; str. 131).

### 1.5.5. Shimmer

*Shimmer* je termin koji se odnosi na intenzitetsku nepravilnost, odnosno fluktuaciju amplitude zvučnog signala. *Shimmer* je brzo kolebanje amplituda i najčešće se izražava u decibelima. Normalna vrijednost *shimmera* dopušta kolebanja do 0.35 dB, a povišene vrijednosti u govornom glasu percipiramo kao promuklost (Heđever, 2012; str. 131 i 132).



**Slika 4.** Prikaz jittera i shimmera na oscilogramu (Preuzeto iz Heđever, 2012).

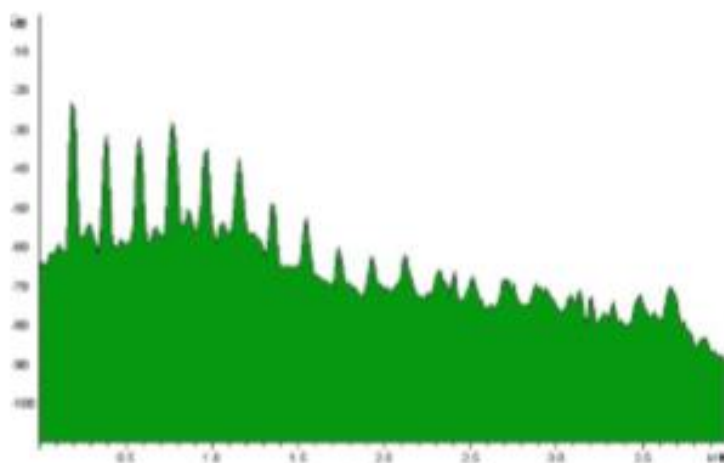
## 1.6. Spektar govora

Spektar je dvodimenzionalan i prikazuje zvučni signal u dvije dimenzije: frekvenciju (*X*-os) i amplitudu (*Y*-os). Spektralna analiza zvuka je postupak kojim se složeni zvučni val razlaže na njegove sastavne frekvencije. Ulazni parametri spektralne analize su amplitude zvuka u vremenu, a izlazni – raspodjela frekvencija u određenom frekvencijskom pojasu u vremenu. Spektralnu analizu možemo najjednostavnije shvatiti tako da se svaki složeni zvuk sastoji od pojedinačnih zvukova različitih frekvencija i amplituda (Heđever, 2012; str. 104 i 105).

### 1.6.1. FFT spektar

Računanje spektralne analize u realnom vremenu jedan je složeni račun, ali razvojem tehnologije, ponajviše računalne omogućeno je korištenje algoritma koji se naziva brza Fourierova transformacija (*Fast Fourier Transform* – FFT). Ime je dobila po francuskom matematičaru i fizičaru Jean Baptiste Joseph Fourieru (1768 – 1830). U svom djelu *Théorie analytique de la chaleur* objavljenom 1822. godine objasnio je da se bilo koja periodična funkcija (u koju spada i zvuk) može približiti, aproksimirati trigonometrijskim polinomom –

Fourierovim redom. Taj se postupak naziva Fourierovom ili harmonijskom (spektralnom) analizom. Maksimalan frekvencijski opseg spektralne analize osnovane na FFT metodi je jednak polovici frekvencije uzorkovanja (*sampling rate*), koja je već prethodno spomenuta u ovom radu i objašnjeno je kolika mora biti frekvencija uzorkovanja i zašto je preporučljivo da bude tolika (Heđever, 2012). Zaključno, Fourierova reprezentacija signala igra važnu ulogu, i u kontinuiranom vremenu procesiranja i u diskretnom vremenu procesiranja signala (Hayes, 1999). Dakle, da bi se neka funkcija mogla razviti u Fourierov red, uvjet je da je funkcija periodična. Fourierova analiza smatra se kao rastavljanje signala na njegove osnovne dijelove – frekvenciju i amplitudu (Lukavečki, 2008). Zaključno, glavni koncept Fourierove analize jest da svaka periodička funkcija može biti zapisana kao suma (ne nužno konačna) sinusa različitih amplituda, faza i frekvencija. Ta suma naziva se Fourierov red (Tolstov, 2014).



**Slika 5.** FFT spektral ljudskog glasa (fonacija vokala /A/) (Preuzeto iz Heđever, 2012).

### 1.6.2. LPC spektral

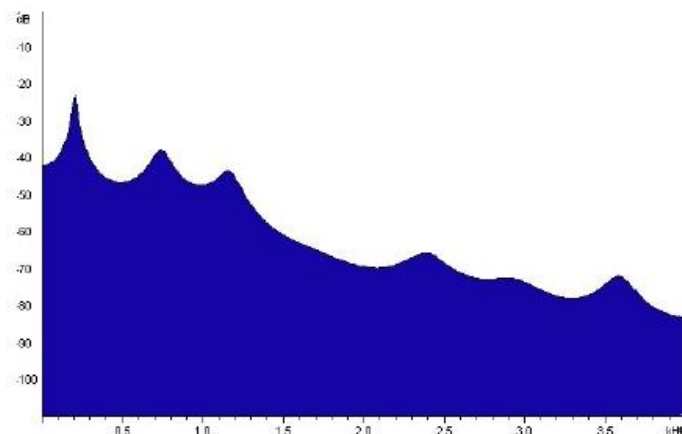
Analiza govora temeljena na linearnoj predikciji (*Linear Predictive Coding* – LPC) najčešće je korištena tehnika za određivanje osnovnih značajki govornog signala (spektral, frekvencije formanata, funkcija vokalnog trakta, modeli i spektralna ovojnica signala, pobudni

signal), kao i za kodiranje s postizanjem visoke kompresije u svrhu prijenosa ili pohrane govornog signala.

Neke od prednosti korištenja LPC-a u analizi govornog signala te u svrhu prepoznavanja govora su:

- *odlična aproksimacija govornog signala, posebice za zvučne glasove*
- *odvajanje modela vokalnog trakta od pobudnog signala te dobivanje spektralne ovojnice*
- *jednostavnost hardverske ili softverske implementacije (Rendulić, 2011; str. 3).*

LPC spektar dobiva se iz FFT analize izračunavanjem prosječnih vrijednosti. Takvim uprosječavanjem FFT spektar postaje „izglačan“, eliminiraju se harmonici (više nisu istaknuti), a veće nakupine pojačanih harmonika tvore formante koji postaju vidljivi (Heđever, 2012; str. 105 i 106).

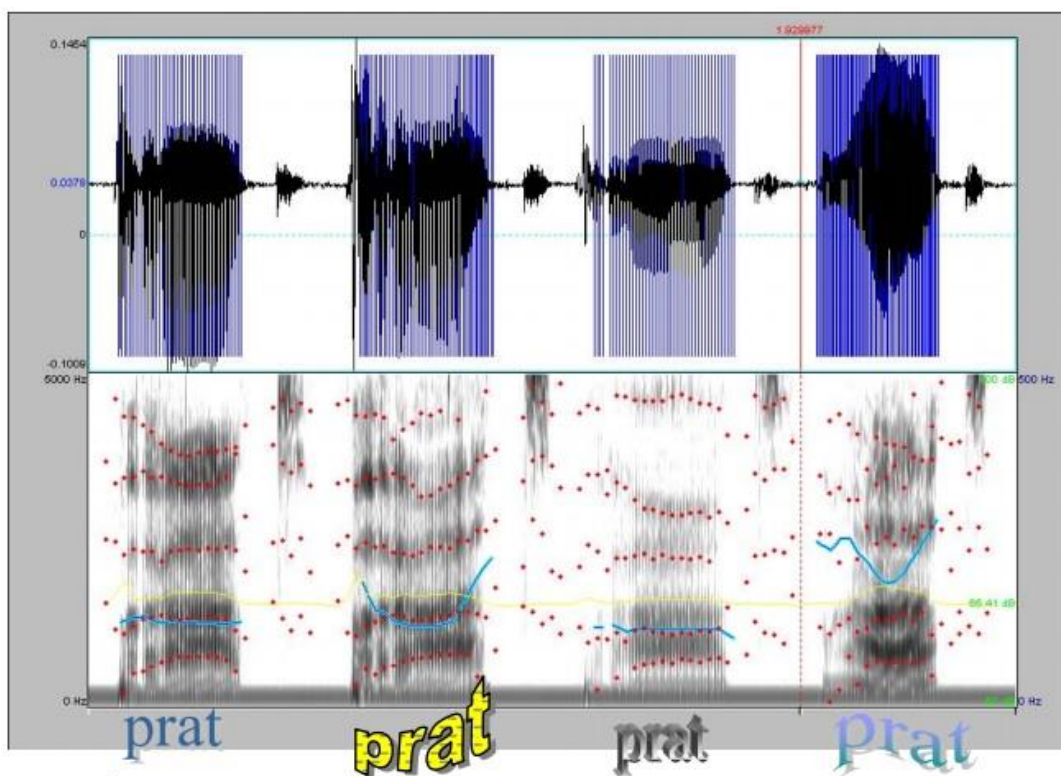


**Slika 6.** LPC spektar ljudskog glasa (fonacija vokala /A/ iz prethodnog primjera).

### 1.6.3. Spektrogram

Naziva se još i sonogram. Za razliku od oscilograma, koji zvuk prikazuje u dvije dimenzije, spektrogram prikazuje zvuk u tri dimenzije :

- vrijeme (x-os)
- frekvencija (y-os)
- amplituda (z-os)



**Slika 7.** Prikaz spektrograma (donji redak) i oscilograma (gornji redak) (Preuzeto iz [https://web.stanford.edu/dept/linguistics/corpora/material/PRAAT\\_workshop\\_manual\\_v421.pdf](https://web.stanford.edu/dept/linguistics/corpora/material/PRAAT_workshop_manual_v421.pdf)).

## 2. CILJ RADA

Nakon što je prethodni dio rada dao detaljniji uvod o raznim pojmovima vezanim uz akustiku, pritom opisujući i relevantne matematičke koncepte, sada će dio tih pojmova biti iskazan u okviru programa PRAAT. Cilj ovoga preglednog rada jest osnovne akustičke parametre prikazati u ovom programu, najprije opisujući osnovne principe rada programa te u koje se sve svrhe program može koristiti. Dodatno će biti opisana funkcija *Praat Script* te čemu služi. U PRAAT-u će se prikazati skript za AVQI (*Acoustic Voice Quality Index*). Na samom kraju bit će opisana primjena ovog programa u logopediji.

### 3. PRAAT

Kompleksni akustički val slušanjem (subjektivnom analizom) nije moguće rastaviti na sastavne dijelove, nego samo akustičkim spektralnim analizama (objektivnom analizom). Tome nam služe različita tehnološka, računalna i softverska rješenja, među kojima je, između ostalih, i program PRAAT.

U kliničkoj praksi najčešće se primjenjuje trodimenzionalna akustička spektralna analiza kojom se, uz primjenu različitih filtara, glas razlaže na sastavne dijelove u određenom vremenskom odsječku (<http://www.foni.mef.hr/Prirucnik/Fonijatrija.htm>).

PRAAT je besplatni računalni softverski paket namijenjen znanstvenoj analizi govora. Koristi se u fonetici, akustici, forenzičkoj akustici i fonetici te logopediji. Autori ovog softverskog paketa su Nizozemci Paul Boersma i David Weenink s Instituta fonetskih znanosti Sveučilišta u Amsterdamu. PRAAT se neprestano revidira i poboljšava novim verzijama. Dostupan je većini operativnih sustava, a neki od njih su Unix, Linux, Macintosh, Solaris i Microsoft Windows.

Iako se PRAAT koristi za složene vrste analiza od strane iskusnih fonetičara, može biti koristan i onima koji imaju manje tehničkog iskustva s ovim programom. Mogu ga koristiti i studenti logopedije koji se upoznaju sa osnovnim obilježjima glasa i govora kod govornika bez patoloških obilježja, ali svakako i o patološkim karakteristikama glasa i govora. Cilj je bio malo više „zagrebati“ po površini ovog programa. Naravno, ostat će sve više-manje samo na osnovama. Zašto? Čovjek općenito najbolje uči sustavno koristeći i vježbajući određenu materiju. Tehnološki napredak u današnje vrijeme doveo je do toga da postoji zaista velik broj istraživanja iz područja akustike. Kao što je već spomenuto, programi namijenjeni analizi glasa i govora neprestano se revidiraju. U moru informacija, izazov je pronaći aktualne skriptove koji se koriste. Potrebno je razumjeti osnove kako bi sve nadogradnje u računalnim programima bile razumljive i korisne u praktičnoj upotrebi. Prema tome, svrha ovog rada jest prikazati osnovne principe rada u programu PRAAT koji je višestruko koristan, kako u znanosti tako i u logopedskoj praksi.

Program nudi široki raspon funkcija i analiza, uključujući spektrogramsku analizu, analizu formantata, analizu parametara glasa (fundamentalna frekvencija, *jitter*, *shimmer* itd.). Sustavno korištenje ovog programa omogućuje kliničarima različitih profila sve veću vještinu

u poznavanju ovog vrijednog softverskog alata, počevši od početnih aktivnosti učenja (koje uključuju i studente logopedije), pa sve do sofisticirane kliničke i istraživačke primjene.



**Slika 8.** Službeni logo programa PRAAT (Preuzeto s <https://grok.lsu.edu/>).

Na tržištu se nudi bezbroj računalnih softverskih programa i paketa za analizu zvuka. Neki od njih su: Adobe Audition, Cool Edit Pro, MDVP (*Multi-Dimensional Voice Program* tvrtke KayPENTAX Corp.), CSL (*Computerized Speech Lab*), Phonanium, a među njima je, naravno, i PRAAT.

PRAAT (nizozemski - *prat* je imperativni oblik glagola *govoriti*). Kao što je već prethodno spomenuto, PRAAT je besplatni softverski računalni paket čija je primarna svrha znanstvena fonetska analiza govora. Kreirali su ga David Weenink i Paul Boersma koji ga zapravo još uvijek ažurno dograđivaju. Primjerice, najnovija izmjena izvršena je 2. svibnja 2019., a autori idu čak toliko detaljno u nadogradnje da čak i objasne što se točno promijenilo ili poboljšalo u najnovijoj verziji. Na službenoj stranici pod opisom najnovije verzije stoji sljedeće: *Removed a bug introduced in 6.0.51 that could cause turning a Discriminant into a Configuration to crash.*", što zaista ide u prilog koliko je autorima stalo do kvalitete ovog softverskog programa, s tim da i sami nude svoj kontakt kako bi im se obratio bilo tko za kakav prijedlog, primjećenu grešku ili konstruktivnu kritiku.

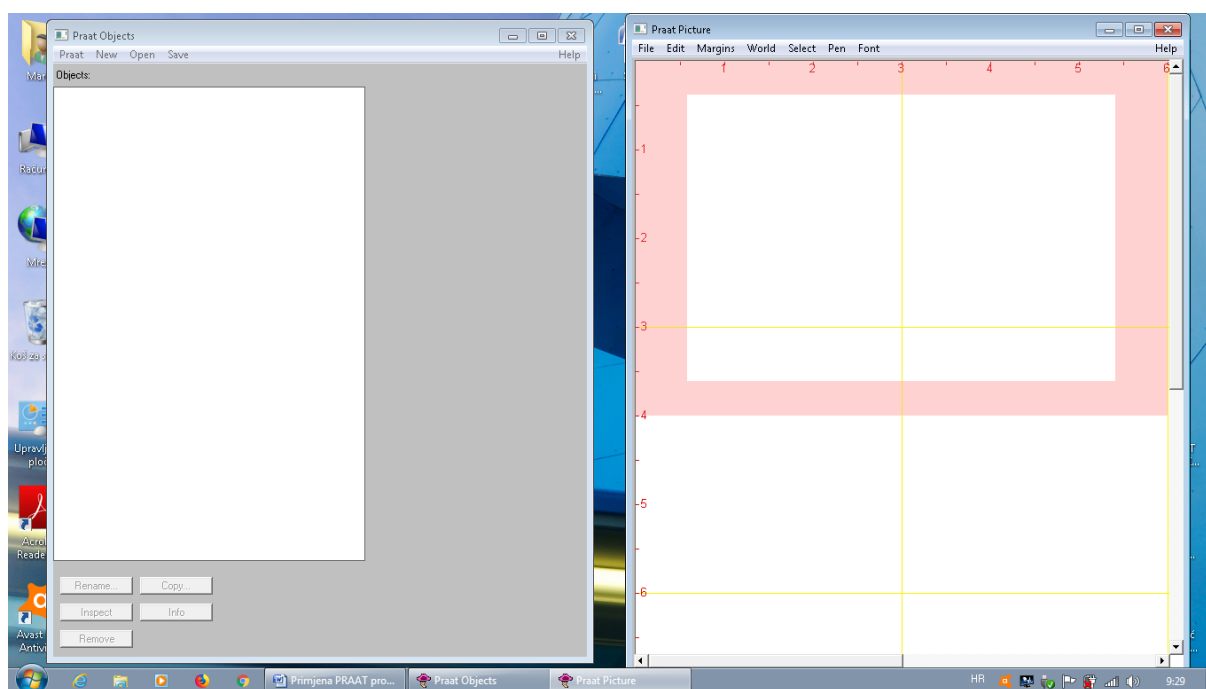
Prva verzija programa (3.1) izrađena je 5. prosinca 1995. godine. Najnovija dostupna verzija (u trenutku pisanja ovoga rada) je 6.0.52 (2. svibnja 2019.) koja se, naravno, besplatno može preuzeti i instalirati sa službene stranice <http://www.fon.hum.uva.nl/praat/>. Program generalno nudi mogućnosti analize i sinteze govora, manipuliranje govorom te bezbroj drugih svojstava, poput spektralne analize, analize fundamentalne frekvencije, analizu formanta, analizu intenziteta, kao i fonetiku i statistiku.

### 3.1. Osnovni principi rada

Ovaj pregledni rad će dati pregled samo osnovnih funkcija programa. Svaka slika zaslona (*screenshot*) učinjena je na vlastitom računalu kako bi upotreba programa bila što vjerodostojnija. Program zaista ima bezbroj mogućnosti i bilo bi nemoguće prikazati apsolutno sve, međutim, takav cilj bio bi i nepotreban. Ako ikad ikome zatreba ovaj program, zaista sve informacije može naći koristeći relevantne internetske izvore, a i sama službena stranica nudi iscrpan uvod i *tutorial*, kao i program u *Help* izborniku.

#### 3.1.1. Pokretanje programa

Prilikom otvaranja, istovremeno se otvaraju dva programska prozora – *Praat Picture* i *Praat Objects*, što prikazuje Slika 9. Snimka zaslona, kao što je prethodno rečeno, napravljena je na vlastitom računalu.

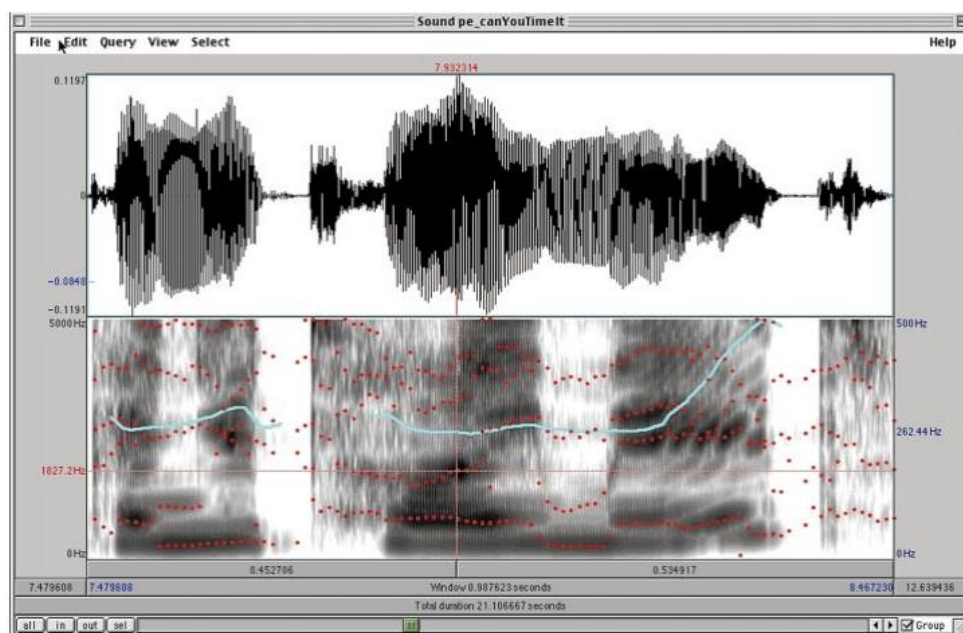


**Slika 9.** Dva prozora pri inicijalnom otvaranju programa PRAAT (*Praat Objects* – lijevo; *Praat Picture* – desno).

Prozor *Praat Objects* ima dalje sljedeće kategorije: ***Praat, New, Open, Save, Praat*** - daje informacije o programu i što uopće sam program jest. Tu se nalaze i dva izbora ***New***

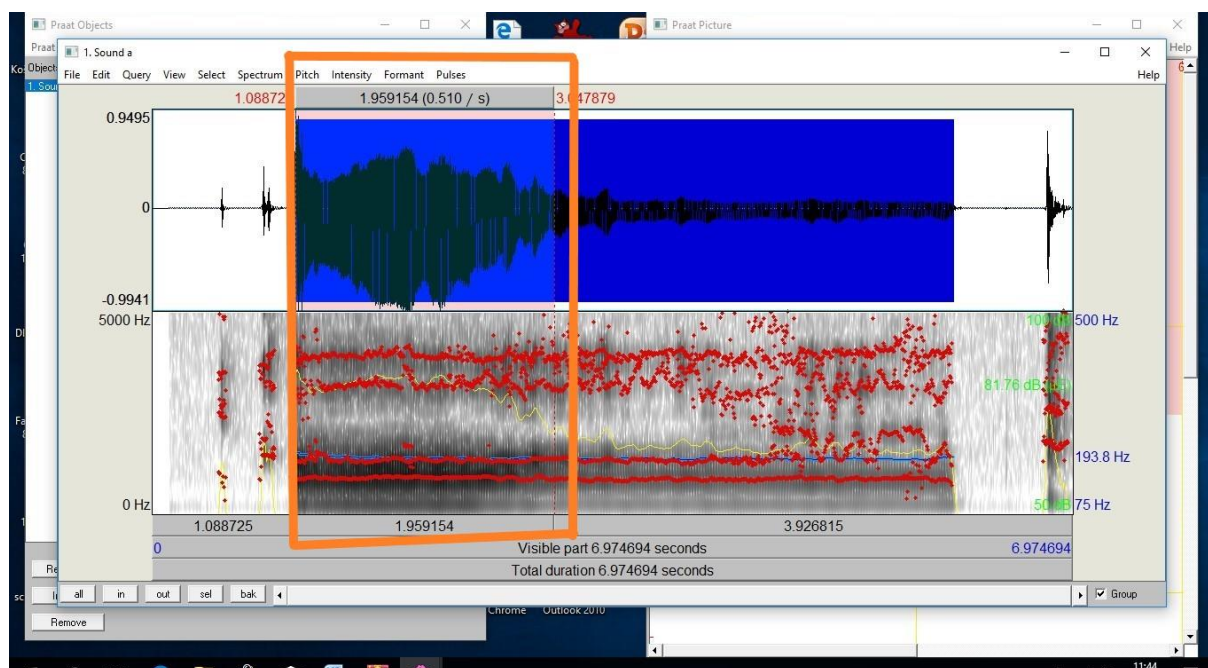
**Praat Script** i **Open Praat script**, o kojima će biti riječ kasnije. Prilikom otvaranja dijela **Open**, najvažnije je istaknuti **Read from file** i **Open long sound file**, što podrazumijeva da na računalu korisnik već ima neke postojeće zvučne zapise koje želi dalje analizirati ili ima one zapise koje je prethodno izradio u samome programu ili je snimio npr. uzorak vlastitog glasa na računalu. Ovdje treba imati na umu kako standardni zvučnik računala ili laptopa implementiran u sam uređaj više služi za osobne svrhe i nikako ne zadovoljava kriterije za kliničke analize, koje povrh toga podrazumijevaju i zadovoljene kriterije akustičke arhitektonike prostora u kojemu se zvuk/glas/govor snima i daljnje analizira. Izbornik **New** daje velik broj radnji koje se mogu izvršiti, a najčešća je **Record mono Sound** koja se kraće može izvršiti istovremenim pritiskom na tipkovnici tipke **Control** plus **R (Ctrl + R)**.

U prozoru **Praat Picture**, ukratko, se prilikom obrade signala izbacuju željeni grafovi (Kovač, 2012). Taj je prozor, također, i izrazito dinamičan. Ovu stavku nemoguće je prikazati u pisanom obliku ovog diplomskog rada jer je to proces koji se događa u vremenu tijekom snimanja uzorka govora. Ono što se može vidjeti tijekom snimanja jesu žuta i zelena boja, a ukoliko glasnoća glasa postane prevelika, pojavljuje se i crvena boja; smanjivanjem glasnoće govora, nestaje i crvena boja.



**Slika 10.** Gornji dio programskog prozora (oscillogram) i donji dio programskog prozora (spektrogram) (Preuzeto iz Boersma i van Heuven, 2001).

### 3.1.2. Analiza zvučnog zapisa



Slika 11. Oscillogram i spektrogram fonacije vokala /A/

Slika 11. Prikazuje fonaciju vokala /A/ ženske 26– godišnje osobe koju sam snimila koristeći PRAAT program instaliran na vlastitom računalu. Kao što je i na slici vidljivo, cijeli zvučni zapis traje 6 sekundi i 97 stotinki. Anotiranjem zapisa (narančasti okvir izrađen alatom Paint 3D) izdvojila sam 1 sekundu i 95 stotinki zapisa jer u tom dijelu nema nikakvih prekida i najbolje vizualno prikazuje fonaciju. Važno je napomenuti da uvjeti snimanja nisu bili idealni najprije zato što je osoba govorila u standardni pripadajući zvučnik na laptopu, a mali radni kabinet nije prostor koji je zvučno izoliran i akustički primjeren. Poanta cijelog zapisa bila je „vlastoručno" na običnom primjeru prikazati kako snimiti neki zapis i vidjeti koji su osnovni parametri kvalitete glasa koji se mogu iščitati uz pomoć PRAAT softverskog alata, odnosno programa. Nakon što sam kliknula nekoliko naredbi u *Praat Objects* prozoru (**New - Record mono Sound**; otvorio se prozor **SoundRecorder**). Iako stereo zvuk omogućava lokalizaciju zvukova u prostoru, generalno fonetska istraživanja funkcioniraju na mono signalima, jer većina mikrofona hvata samo jedan kanal, dokle ljudski vokalni trakt sadrži samo jedan izvor zvuka (Styler, 2017).

Tu treba paziti da se klikne kućica koja nudi *Sampling Frequency* od 44100 Hz (frekvenciju uzorkovanja o kojoj je bilo riječ u prethodnim poglavljima i zašto je preporučljivo da bude baš tolika). Nakon toga klikne se **Record** i snima se. Ovaj je prozor također i izrazito dinamičan. Tu stavku nemoguće je prikazati u pisanom obliku ovog rada jer je to proces koji se događa u realnom vremenu tijekom snimanja uzorka govora. Ono što se može vidjeti tijekom snimanja jesu inicijalna žuta i zelena boja, a ukoliko glasnoća glasa postane prevelika, pojavljuje se i crvena boja; smanjivanjem glasnoće govora, nestaje ta crvena boja. Kad želimo zaustaviti snimanje, klikne se **Stop**. Nakon toga program daje mogućnost proizvoljnog preimenovanja zvučnog zapisa te **Save to List and Close**. Nakon što sam snimila fonaciju vokala /A/, taj isti zapis sam spremila na radnu površinu računala u **.wav** formatu. Taj zapis otvorila sam u programu koristeći **Praat Objects** prozor te u njemu **Open**, a zatim **Read from File** koji naravno nudi mogućnost traženja datoteke po mjestu pohrane i nazivu. Kad je datoteka postala vidljiva (označena plavom bojom) u Praat Objects prozoru, s desne strane prozora klikne se naredba **View & Edit** koja zatim otvara novi prozor prikazujući u gornjem retku oscilogram, a u donjem retku spektrogram zapisa. Na tom prozoru daljnje se potraži u gornjem izborniku prozora naredba **Pulses** te zatim **Voice Report**. U Prilogu 1. vidjeti kako izgleda kompletna lista parametara koje PRAAT prikaže, na primjeru fonacije vokala /A/ iz teksta. Najvažnija tri parametra su *jitter*, *shimmer* i HNR omjer. Parametar HNR (*Harmonics to Noise Ratio*) često se mjeri u fonaciji i pokazuje kolika je razlika između prosječne vrijednosti harmonijskog dijela spektra i razine šuma. Izražava se kao vrijednosna razlika izražena u decibelima. Omjer je sljedeći – što je razlika veća, to je glas kvalitetniji i čišći. Smatra se da razlika u normalnom glasu treba iznositi najmanje 10 do 12 dB (Heðever, 2012). Što je broj veći od 12, glas je bolji. Vrijednost ispod 10 dB indicira da u glasu ima puno šuma i da, ukoliko je potrebno, provjeriti radi li se o značajnom poremećaju glasa, uvažavajući pritom ostale relevantne parametre i cjelokupnu kliničku sliku. PRAAT program temeljem **Voice Report**-a prikazuje sljedeće (parametara ima više (pogledati u Prilogu 1., a izvučeni su oni najrelevantniji):

*Median pitch:* 193.320 Hz (aritmetička sredina je najbolji pokazatelj)

*Minimum pitch:* 189.239 Hz

*Maximum pitch:* 204.554 Hz

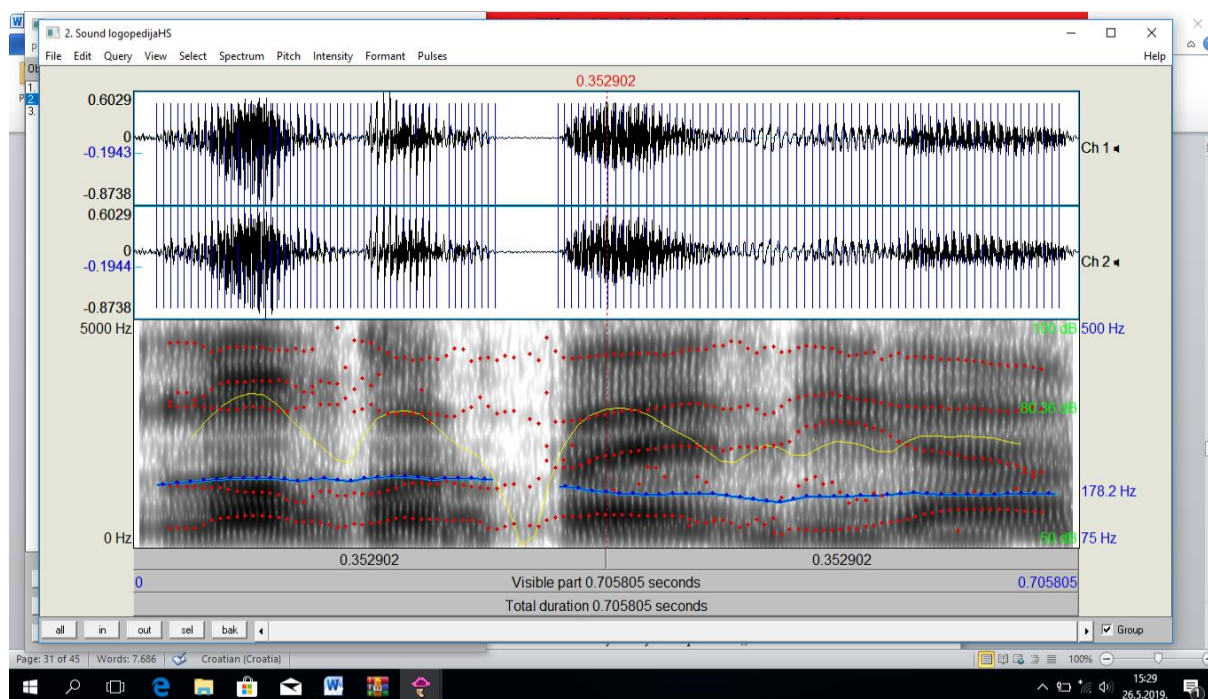
*Jitter (local):* 0.458 %

*Shimmer* (local, dB): 0.473 dB

*Mean harmonics-to-noise ratio*: 19.711 dB

Iz dobivenog može se zaključiti kako postoji povišena vrijednost *shimmer*-a, što nije čudno obzirom da je osoba iz primjera u vrijeme snimanja ovog zapisa bila izrazito prehladna i lagano promukla. Povišenu vrijednost *shimmer*-a u svakom slučaju percipiramo kao promuklost, a povišenu vrijednost *jitter*-a percipiramo kao glas lošije kvalitete, odnosno manje kvalitetan glas kojemu intonacija nije baš „čista“. Također treba napomenuti kako je poželjno da su parametri *Minimum pitch* i *Maximum pitch*, odnosno njihove vrijednosti što bliže vrijednosti fundamentalne frekvencije (*Median pitch*).

Da ne bi sve ostalo na fonaciji i prikazu samo vokala, Slika 12. prikazuje oscilogram i spektrogram izgovorene riječi /logopedija/ koju je izgovorila ženska 30-godišnja osoba u trajanju manjem od jedne sekunde.



**Slika 12.** Riječ /logopedija/ u programu PRAAT (oscilogram i spektrogram).

## 3.2. PRAAT Script

Skriptovi su specijalizirani programski jezici koji služe uštedi vremena na način da automatski generiraju brojne sekvence neke operacije koja se želi izvršiti. Sastoje se od teksta koji sadrži izbornik naredbi. Kad se skript pokrene, naredbe se automatski pokrenu i generiraju. Time se uštedi vrijeme i nije potrebno željenu naredbu zadavati korak po korak. Skoro sve što se može učiniti u PRAAT-u može postati automatizirano koristeći PRAAT Script funkciju. Važno je napomenuti da PRAAT Scripting nije programiranje, već skriptiranje. Kako bi što jednostavnije objasnio skripting jezik programa PRAAT, Boersma kao jedan od njegovih autora, u svom članku iz 2001. navodi sljedeću humorističnu usporedbu: „U većini dijelova svijeta, ropstvo je odbačeno u 19. stoljeću, ali nije neuobičajeno vidjeti fonetičare kako vrijednosti fundamentalne frekvencije 1500 samoglasnika mjere „ručno“. Zasigurno biste htjeli takav rad zamijeniti automatskom procedurom koja prelazi kroz sve zvučne zapise pohranjene u određenom direktoriju ili preko svih segmenata s oznakom "u" u anotaciji samoglasnika (Boersma, 2001; str. 344). Ovakve stvari lako se mogu primijeniti koristeći PRAAT skripting jezik. Mnogi korisnici programa koriste ovaj jezik za analize, tabulacije, statistike (omogućen je čak i Hi-kvadrat test, kao i F testovi). Što se skriptova općenito tiče, svaki korisnik ovog programa tko je malo umješniji služeći se njime, može sam izraditi neke skriptide. Međutim, vrlo korisna stvar koja već postoji u samom programu jest već spomenuti **Praat Objects** prozor koji u izborniku *Help* nudi **Praat Intro** i **Scripting**, a internetska stranica <http://www.fon.hum.uva.nl/praat/manual/Scripting.html> nudi poveći broj već gotovih skriptova. Najveći doprinos u domeni skriptova dao je Youri Maryn, profesor na Fakultetu za medicinu i zdravstvene studije Sveučilišta u Ghentu, Belgija. Autor je i softverskog paketa naziva *Phonanium* (<https://www.phonanium.com/>). U nastavku ovog rada bit će prikazan skript čiji je autor upravo profesor Maryn. Radi se o AVQI skriptu koji je posebno dizajniran za PRAAT program. Neće se ulaziti u tehnički opis niti programski jezik. To nije ni svrha ovoga dijela rada već će se opisno dati uvid u parametar dobiven skriptom te koja je svrha dobivenoga u kontekstu logopedije.

### 3.2.1. AVQI SCRIPT

Uz perceptivnu procjenu kvalitete glasa neizostavno je primijeniti i akustičku analizu glasa, jer njihov komplementaran odnos povećava preciznost u utvrđivanju vokalne disfunkcije te olakšava terapiju istovremeno dokumentirajući njen napredak. „*Glavna karakteristika akustičke analize je upravo njena objektivnost, a perceptivne procjene glasa racionalnost u smislu njezine interpretacije, kao i daljnjeg odabira akustičkih mjernih instrumenata koji će najpreciznije potvrditi ili odbaciti perceptivnu sumnju na poremećaj glasa i omogućiti evaluaciju uspješnosti terapije*” (Bonetti, 2011; str.69).

U području poremećaja glasa akustička analiza igra neizostavnu ulogu, što je sasvim razumljivo jer ona pruža objektivne parametre kvalitete glasa, čineći je objektivnom tehnikom procjene. Savršeno se nadopunjuje sa perceptivnom procjenom glasa. Neke objektivne metode procjene u sebi uključuju analizu više akustičkih parametara, a rezultat je posljedično kombinacija dobivenih parametara. U takvu metodu spada i AVQI (*Acoustic Voice Quality Index*). Autor mu je već spomenuti doaje u području poremećaja glasa, profesor Youri Maryn koji ga je kreirao 2009. godine sa svojim suradnicima. AVQI je dizajniran kako bi mjerio ono što mu i samo ime kaže – izmjeriti generalnu kvalitetu glasa koristeći objektivne i akustičke markere u kliničke svrhe (Maryn i Roy, 2012). AVQI je multivarijantni algoritam koji kombinira više akustičkih markera kako bi se dobio broj koji znatno korelira s generalnom jakosti disfonije. Ovakav „multiparametarski” pristup zapravo proizlazi iz „multidimenzionalne” prirode kvalitete glasa i činjenice da nije povezan s jednom fizikalnom varijablom ili jednim psihoakustičkim izvorom (Núñez-Batalla i sur., 2016). Već je i prihvaćeno da su multiparametarski modeli prikladniji za evaluaciju kvalitete glasa od pojedinačnih mjera. Korištenje višestrukih akustičkih parametara u modelima za procjenu glasa pokazalo je veću korelaciju s auditorno-perceptualnom procjenom te veću pouzdanost i valjanost u detekciji patologije glasa (Uloza i sur., 2018).

Nekoliko subjektivnih skala za procjenu kvalitete glasa vrlo dobro proporcionalno korelira s AVQI, poput GRBAS skale (*Grade, Roughness, Breathiness, Asthenia, Strain*) i CAPE-V skale (*Consensus auditory – perceptual evaluation of voice*). Viši rezultat na subjektivnim skalama korespondira s višim AVQI bodovima, i obrnuto (Maryn i sur., 2010b).

Ono što AVQI posebno izdvaja među alatima za objektivnu akustičku analizu procjene glasa jest to da je to prva metoda koja u sebi sadrži analizu fonacije vokala i kontinuiranoga govora. Procedura je sljedeća: ispitanik treba 3 sekunde fonirati vokal /A/, a zatim naglas pročitati dvije zadane rečenice (Maryn, De Bodt, Barsties i Roy, 2014). Zvučni zapis govora označuje se nastavkom 'cs' (*continuous speech*), a kontinuirana fonacija samoglasnika oznakom 'sv' (*sustained vowel*).

Kako bi izračunao stupanj jakosti disfonije, AVQI koristi šest objektivnih akustičkih parametara, a to su sljedeći:

- lokalni shimmer (*shimmer local*)
- lokalni shimmer u decibelima (*shimmer local dB*)
- HNR omjer
- generalni nagib spektra (*Slope*)
- nagib regresijske linije kroz spektar (*Tilt*)
- zaglađeni kepsralni vrhovi (*CPPS*)

Od navedenih šest parametara, potrebno je istaknuti CPPS parametar – *Cepstral peak prominence smoothed*. Pokazalo se da je ovaj parametar najznačajniji, između ostaloga zato što je baziran na činjenici da što je periodičniji vokalni signal, bolja je definicija harmoničke konfiguracije spektra, što posljedično daje istaknutiji kepsralni vrh. Ovaj parametar, dakle, nudi vrijednu informaciju o stupnju do kojeg dobro definirana harmonijska struktura može biti detektirana u vokalnom signalu, bez da se temelji na detekciji periodičnosti. CPPS je stoga bolji prediktor disfonije nego *jitter* ili *shimmer* (Núñez-Batalla i sur., 2016).

Analizom navedenih šest parametara dobiva se precizan rezultat AVQI-a, odnosno broj na skali od 0 do 10. Klinička granična vrijednost između normofonije i disfonije nalazi se negdje oko 3, ovisno o jeziku i govornom materijalu na kojem je testiran. Rezultat u intervalu od 0 do 3 (praga) označava normofoniju, dok rezultat između praga i 10 označava disfoniju, koja se može razložiti na blagu disfoniju (neposredno iznad praga) do jake disfonije prema broju 10 (Maryn i Weenink, 2015).

Zašto je kontinuirani govor važan? Pokazalo se da sama fonacija vokala nije dovoljna da bi se moglo sa sigurnošću reći o kakvoj se vrsti disfonije radi. Tek se uključivanjem kontinuiranog govora dobila šira slika o kvaliteti ispitivanog glasa. Primjerice, prozodija, ritam, eventualni prekidi glasa i modulacije, kao i dijalekt osobe izraženi su jedino u

kontinuiranom govoru i zaista mogu odlučivati o krajnjoj procjeni i vrednovanju kvalitete glasa (Maryn, De Bodt i Roy, 2010a).

Kvaliteta AVQI-a pokazala se i u kliničkoj i u istraživačkoj uporabi. Pruža relativno brzu i pouzdanu procjenu kvalitete glasa, kao i praćenje svih promjena tijekom terapije, ali i njezine uspješnosti. Za istraživačku uporabu potrebno je normirati kvalitetu glasa u nekom jeziku. Hrvatski jezik još nema svoje ustanovljene norme.

Već su učinjene brojne studije koje su procjenjivale različite aspekte valjanosti AVQI-a u različitim studijama. Prvi aspekt je podudarna valjanost u više istraživanja, što za sobom povlači pitanje koliko dobro AVQI može mjeriti jakost disfonije/promuklosti kroz nekoliko kliničkih uzoraka glasa. Drugi aspekt je dijagnostička preciznost, iz čega pak proizlazi pitanje koliko je AVQI točan u određivanju ima li netko ili nema disfoniju/promuklost. Treći i najvažniji aspekt je reakcija na promjene. Iz toga proizlazi pitanje koliko je AVQI vješt u praćenju auditorno percipiranih promjena u cjelokupnoj kvaliteti glasa tijekom vremena i terapije. Jedan ili više ovih aspekata istraženo je u sljedećim neovisnim studijama: Maryn i sur. (2010a), Maryn i sur. (2010b), Barsties i Maryn (2012), Reynolds i sur. (2012), Maryn i sur. (2014), Barsties i Maryn (2015), Maryn i sur. (2015), Kankare i sur. (2016), Hosokawa i sur. (2016) i Uloza i sur. (2016). Sva ova istraživanja imala su različite sustave snimanja, protokole i okruženja. Njihove studijske skupine sastojale su se od različitih govornika u pogledu demografije i laringealne patologije. Njihovi uzorci kontinuiranog govora potjecali su iz različitih jezika i fonetskih konfiguracija. Sve u svemu, u pouzdanosti mjerenja, AVQI pokazao se kao iznimno robustan. Prosječna korelacija sa subjektivnim skalama iznosi 0,84, a prosječna osjetljivost iznosi 0,85, što ga neminovno čini pouzdanim mjernim instrumentom u kliničkoj procjeni glasa, kao i u terapiji (Maryn i sur., 2014). Tablica 3. prikazuje granične vrijednosti AVQI-a (*cut-off scores*) za jezike u kojima je AVQI već dosad ispitan.

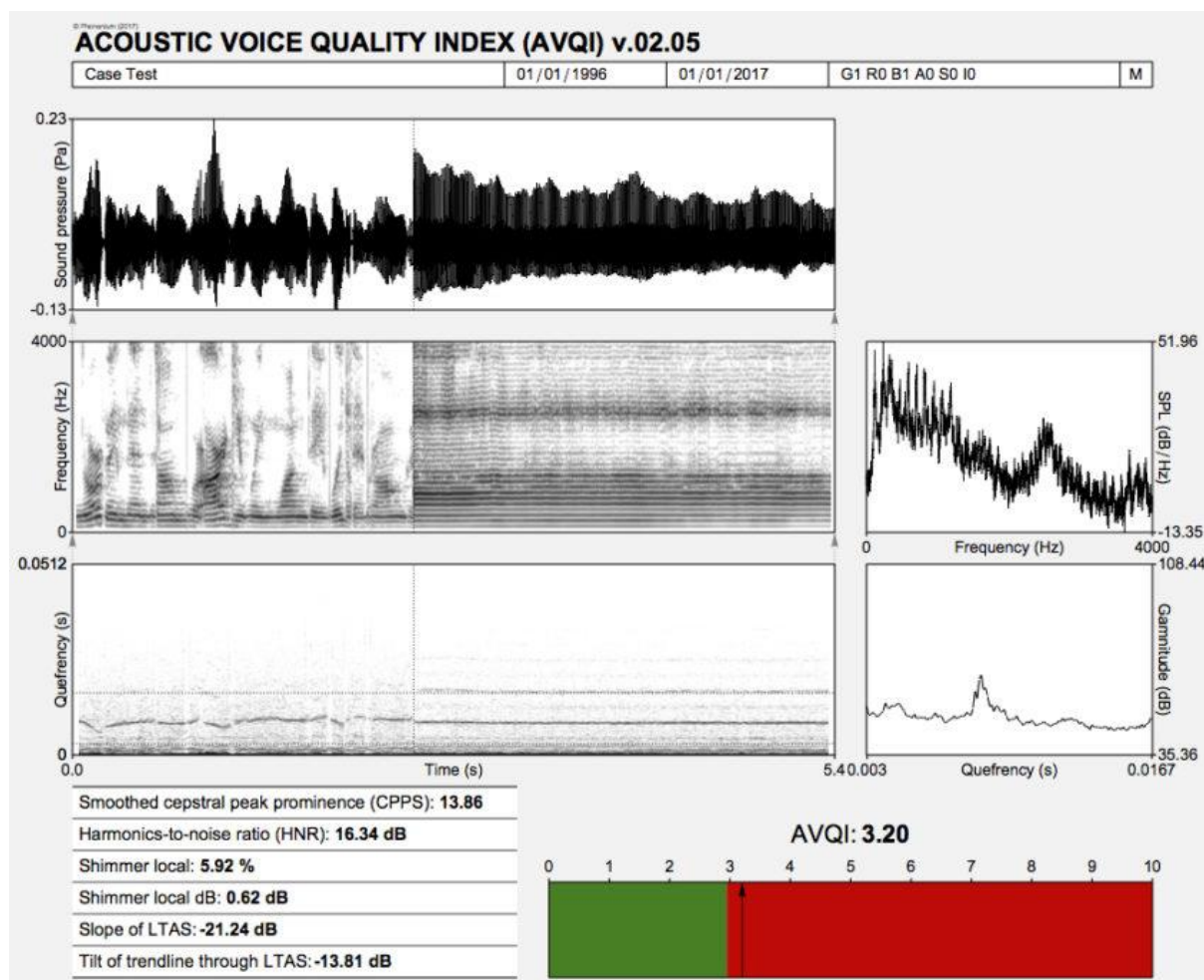
**Tablica 3.** AVQI granične vrijednosti u dosad ispitanim jezicima (Preuzeto iz <https://www.phonanium.com/product/acoustic-voice-quality-index/> ).

| Language            | Threshold scores | Study                   |
|---------------------|------------------|-------------------------|
| Dutch               | 2.95             | Maryn et al. (2010a)    |
| English, Australian | 3.46             | Reynolds et al. (2012)  |
| Finnish             | 3.09             | Kankare et al. (2016)   |
| French              | 3.07             | Maryn et al. (2014)     |
| German              | 2.70             | Barsties & Maryn (2012) |
| Japanese            | 3.15             | Hosokawa et al. (2016)  |
| Korean              | 2.02             | Maryn et al. (2015)     |
| Lithuanian          | 2.97             | Uloza et al. (2016)     |

Slika 13. prikazuje dobivenu vrijednost AVQI-a koristeći gotovi script za AVQI u programu PRAAT. U Prilogu 2. može se vidjeti kako zapravo izgleda script izlistan u cijelosti. Ima puno redaka, s tim da svaki znak # (*hashtag*) označava svojevrsan komentar ili proceduru, a izrazi poput *Remove* i *endif* označavaju naredbu, odnosno, rečeno programskim jezikom, komandu. Rezultat AVQI-a na slici koji iznosi 3.20 označava da se radi o laganoj disfoniji, odnosno poremećaju glasa jer sve što prelazi u crvenu zonu upućuje na patologiju.

Script se u PRAAT-u otvori na sljedeći način: u prozoru **Praat Objects** kliknuti na **Praat**, zatim **Open Praat script...** koji zatim otvara prozor gdje od korisnika traži da otvori folder u kojemu je pohranjen script u tekstualnom obliku. Upravo taj kompletni tekstualni oblik je slijed niz komentara i naredbi koje program PRAAT prepoznaje. Preduvjet je da su prethodno napravljene i pohranjene snimke dva uzorka govora ispitanika, odnosno prethodna fonacija vokala /A/ i uzorak kontinuiranog govora. Ubacujući ta dva zvučna zapisa u program

te onda uključivanjem scripta koji veoma brzo izgenerira ciljane zapise i zatim izbaci broj koji označava vrijednost AVQI indeksa.



**Slika 13.** Acoustic Voice Quality Index dobiven korištenjem istoimenog scripta u programu PRAAT (Preuzeto s <https://www.phonanium.com/product/acoustic-voice-quality-index/>).

#### 4. PRAAT U LOGOPEDIJI

Na Edukacijsko-rehabilitacijskom fakultetu Sveučilišta u Zagrebu izvodi se nekoliko kolegija iz područja akustike i fonetike, kao što su: *Osnove fiziološke i govorne akustike*, *Govorna akustika*, *Psihoakustika* te *Forenzična akustika i fonetika*. Prvi doticaj s ovim zanimljivim područjem i priliku za učenjem imaju svi studenti logopedije. Nadalje, uz MDVP

program, PRAAT čini zlatni standard u identifikaciji govornika u forenzičke svrhe (Hollien, 2002). MDVP i PRAAT su izvanredni alati u akustičkoj i objektivnoj procjeni glasa u raznolike svrhe. Program PRAAT najčešće koriste stručnjaci u području poremećaja glasa, što naravno uključuje logopede koji su se dodatno izučili i za problematiku poremećaja glasa, kao i kako koristiti relevantne alate za analiziranje glasa. PRAAT je najprije istraživački vrlo koristan alat, a zatim i dijagnostički, pogotovo u kombinaciji sa scriptovima. To se moglo vidjeti u prethodnom primjeru na AVQI scriptu, koji analizirajući šest akustičkih parametara daje preciznu vrijednost samog AVQI-a. Ono što bi se inače radilo ručno i zahtijevalo puno više vremena, sada sam program vrlo precizno i u kratkom vremenu izračuna sam. Dobiveni rezultat relevantan je i pouzdan. Primjena samog programa obuhvaćala je najprije poznavanje osnovnih akustičkih pojmova kako bi se znalo uopće što tražiti u samom programu.

## 5. ZAKLJUČAK

Logopedija je interdisciplinarna znanost. Odlikuju je brojna područja kojima se bavi u istraživačke i dijagnostičke svrhe, a jedno od užih područja jest područje akustike. U području logopedije fokus interesa je ljudski govor. Ljudski govor moguće je analizirati subjektivnim, perceptivnim i akustičkim metodama. Ono što je temeljni stup čitave akustike zasigurno jest mjerljivost i objektivnost akustičkih parametara, pa prema tome je potrebno koristiti objektivne i akustičke metode za procjenu kvalitete glasa. Program PRAAT jedan je od brojnih postojećih softverskih paketa. Međutim, ono što ga razlikuje od većine ostalih jest što istovremeno u sebi sadržava besplatnost, ali i kvalitetu koja je priznata od akademske zajednice. Neprestano se revidira i poboljšava novim verzijama koje je moguće preuzeti potpuno besplatno. Štoviše, ne samo da je i jedan od najkompletnijih programa za akustičku analizu, već algoritam zaslužan za analiziranje fundamentalne frekvencije je najprecizniji od svih algoritama koji se trenutno nude za procjenu te stavke (Boersma i van Heuven, 2001). Čini jedan od najvažnijih alata u istraživačke i dijagnostičke svrhe, posebno u kombinaciji sa scriptovima koji su specijalizirani programski jezici koji služe uštedi vremena automatski generirajući brojne sekvencije operacije koja se želi izvršiti. AVQI script je dijagnostički vrlo koristan script, posebno u logopediji, točnije u području poremećaja glasa, iskusnim stručnjacima i kliničarima specijaliziranim za to područje. Laboratorij prepun robustne i brojne mašinerije sada je zamijenjen s jednim računalom ili radnom stanicom. Studenti koji su zainteresirani za područje akustike sada mogu snimiti i analizirati bilo koje vrste zvučnih zapisa, anotirati i modificirati sa nekoliko jednostavnih naredbi ili s nekoliko klikova mišem. Čak i kalkulacija nekih govornih parametara, poput osnovne fundamentalne frekvencije i spektralne analize, koji su u prošlosti bili izuzetno složeni za odrediti, sada se i oni rješavaju s nekoliko klikova mišem. Uloženo vrijeme i trud u poznavanje PRAAT programa daje dugotrajne dobrobiti kliničarima, znanstvenicima, a naposljetku i studentima logopedije koji pokazuju veće zanimanje u područjima akustike i fonetike.

## 6. LITERATURA

**Barsties, B., i Maryn, Y.** (2012). Der Acoustic Voice Quality Index in Deutsch. *Hno*, 60(8), 715-720.

**Beer, M.** (2008). Mathematics and Music: Relating Science to Arts? *Mathematical Spectrum*, 41(1), 36-42.

**Boersma, P. i van Heuven, V.** (2001): Speak and unSpeak with PRAAT. *Glott International*, 5(9), 341-347.

**Bonetti, A.** (2011). Perceptivna procjena glasa. *Hrvatska revija za rehabilitacijska istraživanja*, 47(1), 64-71.

**Culek, I.** (2015). *Glazba i matematika*. Diplomski rad. Osijek: Odjel za matematiku Sveučilišta Josipa Jurja Strossmayera.

**Frosch, H.N.** (2017). *Spectral Analysis and Comparison of Analog and Digital Recordings*. Honors Theses, Honors College, The University of Southern Mississippi.

**Hayes, M. H.** (1999). *Schaum's outline of theory and problems of digital signal processing*. McGraw-Hill.

**Hedeveer, M.** (2012). Osnove fiziološke i govorne akustike. Skripta iz kolegija Osnove fiziološke i govorne akustike na Edukacijsko-rehabilitacijskom fakultetu Sveučilišta u Zagrebu.

**Hedge, M. N.** (2006). *A Coursebook on Aphasia and Other Neurogenic Language Disorders*. New York: Thomson Delam Learning.

**Hollien, H.F.** (2002): *Forensic Voice Identification*. London: Academic Press.

**Hosokawa, K., Barsties, B., Iwahashi, T., Iwahashi, M., Kato, C., Iwaki, S. i Ogawa, M.** (2017). Validation of the acoustic voice quality index in the Japanese language. *Journal of Voice*, 31(2), 260.e1-260.e9.

**Kankare, E., Barsties, B., Maryn, Y., Ilomäki, I., Laukkanen, A. M., Tyrmi, J. i Vilpas, S.** (2015). A preliminary study of the Acoustic Voice Quality Index in Finnish speaking population. In 11th Pan European Voice Conference; 2015 August 31th—September 4th; Florence (p. 218). Firenze University Press, Italy.

**Kovač, Z.** (2012). Prepoznavanje osobe u forenzici na temelju glasa. *Policija i sigurnost*, 21(2), 348-356.

**Lindh, J.** (2017): *Forensic Comparison of Voices, Speech and Speakers: Tools and Methods in Forensic Phonetics*. Doctoral disertation. Department of Philosophy, Linguistics and Theory of Science University of Gothenburg.

**Lukavečki, M.** (2008). *Fourierova analiza akustičkih signala u realnom vremenu*. Diplomski rad. Zagreb: Prirodoslovno - matematički fakultet Sveučilišta u Zagrebu.

**Maryn, Y., De Bodt, M. i Roy, N.** (2010a). The Acoustic Voice Quality Index: Toward improved treatment outcomes assessment in voice disorders. *Journal of Communication Disorders*, 43, 161–174.

**Maryn, Y., Corthals, P., Van Cauwenberge, P., Roy, N., i De Bodt, M.** (2010b). Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *Journal of Voice*, 24(5), 540-555.

**Maryn, Y., De Bodt, M., Barsties, B. i Roy, N.** (2014). The value of the Acoustic Voice Quality Index as a measure of dysphonia severity in subjects speaking different languages. *European Archives of Oto-Rhino-Laryngology*, 271, 1609–1619.

**Maryn, Y. i Weenink, D.** (2015). Objective dysphonia measures in the program Praat: smoothed cepstral peak prominence and acoustic voice quality index. *Journal of Voice*, 29(1), 35-43.

**Maryn, Y., Kim, H. T., i Kim, J.** (2016). Auditory-perceptual and acoustic methods in measuring dysphonia severity of Korean speech. *Journal of Voice*, 30(5), 587-594.

**Maryn, Y. i Roy, N.** (2012). Sustained vowels and continuous speech in the auditory-perceptual evaluation of dysphonia severity. *Jornal da sociedade Brasileira de fonoaudiologia*, 24(2), 107-12.

**Núñez-Batalla, F., Díaz-Fresno, E., Álvarez-Fernández, A., Muñoz Cordero, G. i Llorente Pendás, J.L.** (2016). Application of the Acoustic Voice Quality Index for Objective Measurement of Dysphonia Severity. *Acta Otorrinolaringol Esp.* 2017, 68(4), 204-211.

**Milivojević, Z., Milosavljević, D. i Rajković, M.** (2014). Algoritam za procenu inharmoničnosti trozvuka. *INFOTEH – JAHORINA*, 13, 451-456.

**Petrinović, D.** (2002): Digitalna obrada govora. Interna zavodska skripta pri Zavodu za elektroničke sustave i obradbu informacija na Fakultetu elektrotehnike i računarstva Sveučilišta u Zagrebu.

**Pinculić, H.** (2015). *Obrada eksperimentalnih podataka procesa zakaljivanja primjenom osrednjavajućih i interpolacijskih filtara*. Završni rad. Zagreb: Fakultet strojarstva i brodogradnje Sveučilišta u Zagrebu.

**Rendulić, I.** (2011). *Mjere udaljenosti u obradi govornog signala*. Završni rad. Zagreb: Fakultet elektrotehnike i računarstva Sveučilišta u Zagrebu.

**Reynolds, V., Buckland, A., Bailey, J., Lipscombe, J., Nathan, E., Vijayasekaran, S., i French, N.** (2012). Objective assessment of pediatric voice disorders with the acoustic voice quality index. *Journal of Voice*, 26(5), 672.e1-672.e7.

**Stone, M.** (2001). *Mathematics for Physics II*. Alexandria, Florence, London: Pimander-Casaubon.

**Škarić, I.** (1991). Fonetika hrvatskoga književnog jezika. U: *Povijesni pregled, glasovi i oblici hrvatskoga književnog jezika* (Ured.: Katičić, R.), Hrvatska akademija znanosti i umjetnosti i "Globus" nakladni zavod, Zagreb, , 61 - 372.

**Titze, I.R., Talkin, D.** (1979). A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation. *Journal of the Acoustical Society of America*, 66, 60-74.

**Tiwari, M.** (2012). Speech acoustics: How much science?. *Journal of Natural Science, Biology and Medicine*, 3(1), 24-31.

**Tolstov, G.P.** (2014). *Fourier Series*. Translated from the Russian by Richard A. Silverman. New York: Dover Publications Inc.

**Uloza, V., Barsties von Latoszek, B, Petrauskas, T. i Maryn, Y.** (2018). A comparison of Dysphonia Severity Index and Acoustic Voice Quality Index measures in differentiating normal and dysphonic voices. *European Archives of Oto-Rhino-Laryngology*, 275(4), 949-958.

**Uloza V., Petrauskas T., Padervinskis E., Ulozaitė, N., Barsties von Latoszek B. i Maryn, Y.** (2016). Validation of the Acoustic Voice Quality Index in the Lithuanian language. *Journal of Voice*, Jul 14, epub ahead of print.

### **Elektronički izvori:**

AVQI. <https://www.phonanium.com/product/acoustic-voice-quality-index/> Pristupljeno 20. svibnja 2019.

Boersma, P. and Weenink, D. (2012). Praat: doing phonetics by computer (version 6.0.52). Pristupljeno 10. svibnja 2019. <http://www.fon.hum.uva.nl/praat/>

CPLOL (Comité Permanent de Liaison des Orthophonistes / Logopedes de l'Union Européenne). Pristupljeno 12. svibnja 2019. <https://cplol.eu/>

Encyclopaedia Britannica (2019). *Harry Nyquist, American physicist*. Pristupljeno 14. svibnja 2019. <https://www.britannica.com/biography/Harry-Nyquist>

Gans, D. (2015). *Digital vs Analog Audio: An Overview*. Pristupljeno 10. svibnja 2019. <https://www.klipsch.com/blog/digital-vs-analog-audio>

Leksikografski zavod Miroslav Krleža. Analogno-digitalna pretvorba. Pristupljeno 14. svibnja 2019. <http://www.enciklopedija.hr/natuknica.aspx?ID=2475>

Phonanium. <https://www.phonanium.com/> Pristupljeno 20. svibnja 2019.

Proleksis enciklopedija (2012). *Timbar*. Pristupljeno 14. svibnja 2019. <http://proleksis.lzmk.hr/57191/>

Scripting. <http://www.fon.hum.uva.nl/praat/manual/Scripting.html> Pristupljeno 20. svibnja 2019.

Styler, W. (2017). Using Praat for Linguistic Research (version 1.8) Pristupljeno 15. svibnja 2019. <http://wstyler.ucsd.edu/praat/UsingPraatforLinguisticResearchLatest.pdf>

[https://www.fer.unizg.hr/download/repository/PrZv\\_10\\_2018.pdf](https://www.fer.unizg.hr/download/repository/PrZv_10_2018.pdf) Pristupljeno 15. svibnja 2019.

<http://www.foni.mef.hr/Prirucnik/Fonijatrija.htm> Pristupljeno 15. svibnja 2019.

<http://www.erf.unizg.hr/hr/studiji/preddiplomski-logopedija>   Pristupljeno 28. svibnja 2019.

<http://www.erf.unizg.hr/hr/studiji/diplomski-logopedija>   Pristupljeno 28. svibnja 2019.

## POPIS SLIKA I TABLICA

|                                                                                     |    |
|-------------------------------------------------------------------------------------|----|
| <b>Slika 1.</b> Analogni i digitalni signal.....                                    | 4  |
| <b>Slika 2.</b> Općeniti princip digitalne obrade signala.....                      | 6  |
| <b>Slika 3.</b> Valna duljina periodičnoga vala.....                                | 9  |
| <b>Slika 4.</b> Prikaz <i>jittera</i> i <i>shimmera</i> na oscilogramu.....         | 13 |
| <b>Slika 5.</b> FFT spektar ljudskog glasa.....                                     | 14 |
| <b>Slika 6.</b> LPC spektar ljudskog glasa.....                                     | 15 |
| <b>Slika 7.</b> Prikaz spektograma i oscilograma.....                               | 16 |
| <b>Slika 8.</b> Službeni logo programa PRAAT.....                                   | 19 |
| <b>Slika 9.</b> <i>Praat Objects</i> i <i>Praat Picture</i> prozori.....            | 20 |
| <b>Slika 10.</b> Oscilogram i spektogram.....                                       | 21 |
| <b>Slika 11.</b> Oscilogram i spektogram fonacije vokala /A/.....                   | 22 |
| <b>Slika 12.</b> Riječ /logopedija/ u programu PRAAT (oscilogram i spektogram)..... | 24 |
| <b>Slika 13.</b> <i>Acoustic Voice Quality Index</i> u programu PRAAT.....          | 30 |
| <br>                                                                                |    |
| <b>Tablica 1.</b> Frekvencija uzorkovanja i kvaliteta zvuka.....                    | 7  |
| <b>Tablica 2.</b> Alikvotni tonovi iznad tona a.....                                | 11 |
| <b>Tablica 3.</b> AVQI granične vrijednosti u dosad ispitanim jezicima.....         | 29 |

## PRILOG 1.

### Parametri dobiveni korištenjem *Voice Report*-a u programu PRAAT

-- Voice report for 3. Sound a --

Date: Sun May 12 14:46:56 2019

WARNING: some of the following measurements may be imprecise.

For more precision, go to "Pitch settings" and choose "Optimize for voice analysis".

Time range of SELECTION

From 1.075306 to 3.027751 seconds (duration: 1.952444 seconds)

Pitch:

Median pitch: 193.320 Hz

Mean pitch: 193.867 Hz

Standard deviation: 2.882 Hz

Minimum pitch: 189.239 Hz

Maximum pitch: 204.554 Hz

Pulses:

Number of pulses: 376

Number of periods: 375

Mean period: 5.158437E-3 seconds

Standard deviation of period: 0.078227E-3 seconds

Voicing:

Fraction of locally unvoiced frames: 0.513% (1 / 195)

Number of voice breaks: 0

Degree of voice breaks: 0 (0 seconds / 0 seconds)

Jitter:

Jitter (local): 0.458%

Jitter (local, absolute): 23.601E-6 seconds

Jitter (rap): 0.263%

Jitter (ppq5): 0.265%

Jitter (ddp): 0.788%

Shimmer:

Shimmer (local): 5.029%

Shimmer (local, dB): 0.473 dB

Shimmer (apq3): 2.783%

Shimmer (apq5): 2.879%

Shimmer (apq11): 3.298%

Shimmer (dda): 8.349%

Harmonicity of the voiced parts only:

Mean autocorrelation: 0.980837

Mean noise-to-harmonics ratio: 0.023007

Mean harmonics-to-noise ratio: 19.711 dB

## PRILOG 2.

### Script ACOUSTIC VOICE QUALITY INDEX (AVQI), v.02.01

(Autor: Youri Maryn)

# TITLE OF THE SCRIPT: ACOUSTIC VOICE QUALITY INDEX (AVQI), v.02.01

# Form for introduction and/or parameterization

form Acoustic Voice Quality Index, v.02.01

comment >>> It is advocated to estimate dysphonia severity in both

comment continuous speech (i.e., 'cs') and sustained vowel (i.e., 'sv') (Maryn et al.,

comment 2010). This script runs on these two types of recordings, and it

comment is important to name these recordings 'cs' and 'sv', respectively.

#comment >>> This script automatically (a) searches, extracts and then concatenates

#comment the voiced segments of the continuous speech recording to a new sound; (b)

#comment concatenates the sustained vowel recording to the new sound, (c) determines

#comment the Smoothed Cepstral Peak Prominence, the Shimmer Local, the Shimmer

#comment Local dB, the LTAS-slope, the LTAS-tilt and the Harmonics-to-Noise Ratio of

#comment the concatenated sound signal, (d) calculates the AVQI-score mostly based

#comment on the method of Maryn et al. (2010), and draws the oscillogram, the narrow-

#comment band spectrogram with LTAS and the power-cepstrogram with power-

#comment cepstrum of the concatenated sound signal to allow further interpretation.

comment >>> For the AVQI to be reliable, it is imperative that the sound recordings

comment are made in an optimal data acquisition conditions.

comment >>> There are two versions in this script: (1) a simple version (only AVQI with

comment data of acoustic measures), and (2) an illustrated version (AVQI with data of

comment acoustic measures and graphs).

choice version: 1

button simple

button illustrated

comment >>> Additional information (optional):

sentence name\_patient

sentence left\_dates\_(birth\_-\_assessment)

sentence right\_dates\_(birth\_-\_assessment)

comment

comment Script credits: Youri Maryn (PhD) and Paul Corthals (PhD)

endform

Erase all

Select inner viewport... 0.5 7.5 0.5 4.5

Axes... 0 1 0 1

Black

Text special... 0.5 centre 0.6 half Helvetica 12 0 Depending on the duration and/or the sample rate of the

Text special... 0.5 centre 0.4 half Helvetica 12 0 sound files, it takes more or less time to process them.

# -----

# PART 1:

# DETECTION, EXTRACTION AND CONCATENATION OF

# THE VOICED SEGMENTS IN THE RECORDING

# OF CONTINUOUS SPEECH.

# -----

select Sound cs

Copy... original

samplingRate = Get sampling frequency

intermediateSamples = Get sampling period

Create Sound... onlyVoice 0 0.001 'samplingRate' 0

select Sound original

To TextGrid (silences)... 50 0.003 -25 0.1 0.1 silence sounding

select Sound original

plus TextGrid original

Extract intervals where... 1 no "does not contain" silence

Concatenate

select Sound chain

Rename... onlyLoud

globalPower = Get power in air

select TextGrid original

Remove

select Sound onlyLoud

signalEnd = Get end time

```

windowBorderLeft = Get start time

windowWidth = 0.03

windowBorderRight = windowBorderLeft + windowWidth

globalPower = Get power in air

voicelessThreshold = globalPower*(30/100)


select Sound onlyLoud

extremeRight = signalEnd - windowWidth

while windowBorderRight < extremeRight

    Extract part... 'windowBorderLeft' 'windowBorderRight' Rectangular 1.0 no

    select Sound onlyLoud_part

    partialPower = Get power in air

    if partialPower > voicelessThreshold

        call checkZeros 0

        if (zeroCrossingRate <> undefined) and (zeroCrossingRate < 3000)

            select Sound onlyVoice

            plus Sound onlyLoud_part

            Concatenate

            Rename... onlyVoiceNew

            select Sound onlyVoice

            Remove

            select Sound onlyVoiceNew

            Rename... onlyVoice

        endif

    endif

    select Sound onlyLoud_part

    Remove

    windowBorderLeft = windowBorderLeft + 0.03

    windowBorderRight = windowBorderLeft + 0.03

    select Sound onlyLoud

endwhile

select Sound onlyVoice


procedure checkZeros zeroCrossingRate

    start = 0.0025

```

```

startZero = Get nearest zero crossing... 'start'
findStart = startZero
findStartZeroPlusOne = startZero + intermediateSamples
startZeroPlusOne = Get nearest zero crossing... 'findStartZeroPlusOne'
zeroCrossings = 0
strips = 0

while (findStart < 0.0275) and (findStart <> undefined)
    while startZeroPlusOne = findStart
        findStartZeroPlusOne = findStartZeroPlusOne + intermediateSamples
        startZeroPlusOne = Get nearest zero crossing... 'findStartZeroPlusOne'
    endwhile
    afstand = startZeroPlusOne - startZero
    strips = strips +1
    zeroCrossings = zeroCrossings +1
    findStart = startZeroPlusOne
endwhile
zeroCrossingRate = zeroCrossings/afstand
endproc

# -----
# PART 2:
# DETERMINATION OF THE SIX ACOUSTIC MEASURES
# AND CALCULATION OF THE ACOUSTIC VOICE QUALITY INDEX.
# -----

select Sound sv
durationVowel = Get total duration
durationStart=durationVowel-3
if durationVowel>3
    Extract part... durationStart durationVowel rectangular 1 no
    Rename... sv2
elseif durationVowel<=3
    Copy... sv2
endif

```

```

select Sound onlyVoice
durationOnlyVoice = Get total duration
plus Sound sv2
Concatenate
Rename... avqi
durationAll = Get total duration
minimumSPL = Get minimum... 0 0 None
maximumSPL = Get maximum... 0 0 None

# Narrow-band spectrogram and LTAS

To Spectrogram... 0.03 4000 0.002 20 Gaussian
select Sound avqi
To Ltas... 1
minimumSpectrum = Get minimum... 0 4000 None
maximumSpectrum = Get maximum... 0 4000 None

# Power-cepstrogram, Cepstral peak prominence and Smoothed cepstral peak prominence

select Sound avqi
To PowerCepstrogram... 60 0.002 5000 50
cpps = Get CPPS... no 0.01 0.001 60 330 0.05 Parabolic 0.001 0 Straight Robust
To PowerCepstrum (slice)... 0.1
maximumCepstrum = Get peak... 60 330 None

# Slope of the long-term average spectrum

select Sound avqi
To Ltas... 1
slope = Get slope... 0 1000 1000 10000 energy

# Tilt of trendline through the long-term average spectrum

select Ltas avqi
Compute trend line... 1 10000
tilt = Get slope... 0 1000 1000 10000 energy

```

```
# Amplitude perturbation measures
```

```
select Sound avqi
```

```
To PointProcess (periodic, cc)... 50 400
```

```
Rename... avqi1
```

```
select Sound avqi
```

```
plus PointProcess avqi1
```

```
percentShimmer = Get shimmer (local)... 0 0 0.0001 0.02 1.3 1.6
```

```
shim = percentShimmer*100
```

```
shdb = Get shimmer (local_dB)... 0 0 0.0001 0.02 1.3 1.6
```

```
# Harmonic-to-noise ratio
```

```
select Sound avqi
```

```
To Pitch (cc)... 0 75 15 no 0.03 0.45 0.01 0.35 0.14 600
```

```
select Sound avqi
```

```
plus Pitch avqi
```

```
To PointProcess (cc)
```

```
Rename... avqi2
```

```
select Sound avqi
```

```
plus Pitch avqi
```

```
plus PointProcess avqi2
```

```
voiceReport$ = Voice report... 0 0 75 600 1.3 1.6 0.03 0.45
```

```
hnr = extractNumber (voiceReport$, "Mean harmonics-to-noise ratio: ")
```

```
# Calculation of the AVQI
```

```
avqi = ((3.295-(0.111*cpps)-(0.073*hnr)-(0.213*shim)+(2.789*shdb)-(0.032*slope)+(0.077*tilt))*2.208)+1.797
```

```
# oud avqi = (3.295-(0.111*cpps)-(0.073*hnr)-(0.213*shim)+(2.789*shdb)-(0.032*slope)+(0.077*tilt))*2.571
```

```
# -----
```

```
# PART 3:
```

```
# DRAWINGS ALL THE INFORMATION AND THE GRAPHS.
```

```
# -----
```

```
# Title and patient information
```

```

Erase all

Solid line

Line width... 1

Black

Helvetica

Select inner viewport... 0 8 0 0.5

Font size... 1

Select inner viewport... 0.5 7.5 0.1 0.15

Axes... 0 1 0 1

Text... 0 Left 0.5 Half Script: Youri Maryn (PhD) and Paul Corthals (PhD)

Font size... 12

Select inner viewport... 0.5 7.5 0 0.5

Axes... 0 1 0 1

Text... 0 Left 0.5 Half ##ACOUSTIC VOICE QUALITY INDEX (AVQI)#

Font size... 8

Select inner viewport... 0.5 7.5 0 0.5

Axes... 0 1 0 3

Text... 1 Right 2.3 Half %%'name_patient$'%

Text... 1 Right 1.5 Half %%'left_dates$'%

Text... 1 Right 0.7 Half %%'right_dates$'%


# Simple version


if version = 1


    # Data


Font size... 10

Select inner viewport... 0.5 7.5 0.5 2

Axes... 0 7 6 0

Text... 0.05 Left 0.5 Half Smoothed cepstral peak prominence (CPPS): ##'cpps:2'#

Text... 0.05 Left 1.5 Half Harmonics-to-noise ratio: ##'hnr:2' dB#

Text... 0.05 Left 2.5 Half Shimmer local: ##'shim:2' \% #

Text... 0.05 Left 3.5 Half Shimmer local dB: ##'shdb:2' dB#

Text... 0.05 Left 4.5 Half Slope of LTAS: ##'slope:2' dB#

```

Text... 0.05 Left 5.5 Half Tilt of trendline through LTAS: ##tilt:2' dB#

Select inner viewport... 0.5 3.8 0.5 2

Draw inner box

Font size... 8

Arrow size... 1

Select inner viewport... 4 7.5 1.25 2

Axes... 0 10 1 0

Paint rectangle... green 0 2.91 0 1

Paint rectangle... red 2.91 10 0 1

Draw arrow... avqi 1 avqi 0

Draw inner box

Marks top every... 1 1 yes yes no

Font size... 16

Select inner viewport... 4 7.5 0.5 1.15

Axes... 0 1 0 1

Text... 0.5 Centre 0.5 Half AVQI: ##avqi:2'#

# Copy Praat picture

Select inner viewport... 0.5 7.5 0 2

Copy to clipboard

# Illustrated version

elsif version = 2

# Oscillogram

Select inner viewport... 0.5 5 0.5 2.0

select Sound avqi

Draw... 0 0 0 0 no Curve

Draw inner box

One mark left... minimumSPL no yes no 'minimumSPL:2'

One mark left... maximumSPL no yes no 'maximumSPL:2'

Text left... no Sound pressure level (Pa)

One mark bottom... 0 no yes no 0.00

One mark bottom... durationOnlyVoice no no yes  
One mark bottom... durationAll no yes no 'durationAll:2'  
Text bottom... no Time (s)

#### # Narrow-band spectrogram

Select inner viewport... 0.5 5 2.3 3.9  
select Spectrogram avqi  
Paint... 0 0 0 4000 100 yes 50 6 0 no  
Draw inner box  
One mark left... 0 no yes no 0  
One mark left... 4000 no yes no 4000  
Text left... no Frequency (Hz)  
One mark bottom... 0 no yes no 0.00  
One mark bottom... durationOnlyVoice no no yes  
One mark bottom... durationAll no yes no 'durationAll:2'  
Text bottom... no Time (s)

#### # LTAS

Select inner viewport... 5.4 7.5 2.3 3.9  
select Ltas avqi  
Draw... 0 4000 minimumSpectrum maximumSpectrum no Curve  
Draw inner box  
One mark left... minimumSpectrum no yes no 'minimumSpectrum:2'  
One mark left... maximumSpectrum no yes no 'maximumSpectrum:2'  
Text left... no Sound pressure level (dB/Hz)  
One mark bottom... 0 no yes no 0  
One mark bottom... 4000 no yes no 4000  
Text bottom... no Frequency (Hz)

#### # Power-cepstrogram

Select inner viewport... 0.5 5 4.1 5.8  
select PowerCepstrogram avqi  
Paint... 0 0 0.00303 0.01667 0 0 no

Draw inner box

One mark left... 0.00303 no yes no 0.003

One mark left... 0.01667 no yes no 0.017

Text left... no Quefrency (s)

One mark bottom... 0 no yes no 0.00

One mark bottom... durationOnlyVoice no no yes

One mark bottom... durationAll no yes no 'durationAll:2'

Text bottom... no Time (s)

# Power-cepstrum

Select inner viewport... 5.4 7.5 4.1 5.8

select PowerCepstrum avqi\_0\_100

Draw... 0.00303 0.01667 0 0 no

Draw tilt line... 0.00303 0.01667 0 0 0.00303 0.01667 Straight Robust

Draw inner box

One mark left... maximumCepstrum no yes no 'maximumCepstrum:2'

Text left... no Amplitude (dB)

One mark bottom... 0.00303 no yes no 0.003

One mark bottom... 0.01667 no yes no 0.017

Text bottom... no Quefrency (s)

# Data

Font size... 10

Select inner viewport... 0.5 7.5 6 7.4

Axes... 0 7 6 0

Text... 0.05 Left 0.5 Half Smoothed cepstral peak prominence (CPPS): ##'cpps:2' #

Text... 0.05 Left 1.5 Half Harmonics-to-noise ratio: ##'hnr:2' dB#

Text... 0.05 Left 2.5 Half Shimmer local: ##'shim:2' \% #

Text... 0.05 Left 3.5 Half Shimmer local dB: ##'shdb:2' dB#

Text... 0.05 Left 4.5 Half Slope of LTAS: ##'slope:2' dB#

Text... 0.05 Left 5.5 Half Tilt of trendline through LTAS: ##'tilt:2' dB#

Select inner viewport... 0.5 3.8 6 7.4

Draw inner box

Font size... 8

Arrow size... 1  
Select inner viewport... 4 7.5 6.75 7.4  
Axes... 0 10 1 0  
Paint rectangle... green 0 2.91 0 1  
Paint rectangle... red 2.91 10 0 1  
Draw arrow... avqi 1 avqi 0  
Draw inner box  
Marks top every... 1 1 yes yes no  
Font size... 16  
Select inner viewport... 4 7.5 5.9 6.65  
Axes... 0 1 0 1  
Text... 0.5 Centre 0.5 Half AVQI: ##'avqi:2'#

# Copy Praat picture

Select inner viewport... 0.5 7.5 0 7.4  
Copy to clipboard

endif

# Remove intermediate objects

select all  
minus Sound cs  
minus Sound sv  
Remove