

Speech Processing Assignment

ABHISHEK KUMAR

18MCMC10

Question 1: List various speech processing applications in-terms of text to speech System. Give proper reference of those listed application.

Ans: **Applications** : Text-To-Speech System (TTS) is a computer-based system that automatically converts text into artificial human speech [1]. Text-To-Speech synthesizers do not playback recorded speech; rather, they generate sentences using plain text as input. It is necessary to distinguish between Text-To-Speech synthesizers from Voice Response Systems.

Voice Response Systems simply concatenate words and segments of sentences and are applicable only in situations where limited vocabulary is required, and pronunciation restrictions exist. Speech synthesizers being considered in this context actually encapsulate models of the human vocal tract to produce a synthetic human voice output corresponding to input text. Since it is impracticable to store pre-recorded audio clips of all words of a language, automatic 'pronunciation' of words and sentences is necessary in TextTo-Speech systems.

1. Applications For The Blind
2. Applications For The Deafened And Vocally Handicapped
3. Educational Applications
4. Applications For Telecommunications And Multimedia

Reference:

1. http://research.spa.aalto.fi/publications/theses/lemmetty_mst/chap6.html#:~:text=Synthetic%20speech%20may%20be%20used,applications%2C%20suc%20as%20talking%20heads.&text=For%20example%2C%20better%20availability%20of,for%20people%20with%20communication%20difficulties.
2. <https://core.ac.uk/download/pdf/83592918.pdf>

Question 2: List out various Text to Speech systems available that are working for Indian language. Give appropriate references for each listed item.

Ans: TTS system integrated with screen reader would be a potential assistive technology which will help visually challenged section and learning disabled section of the society to use the benefits of ICT and knowledge sharing.

Application development under TTS Project:

1. **TTS integrated with Screen Reader for Visually Challenged persons :** TTS integrated with Screen Reader are available in Hindi, Bengali, Marathi, Tamil, Telugu and Malayalam.
2. **Browser Plug-in:** TTS as a browser plug-ins are also developed for eight Indian Languages namely Hindi, Bengali, Marathi, Tamil, Telugu, Malayalam, Odia and Gujarati.
3. **SMS Reader in Indian Languages - Sandesh Pathak.**
4. **Website to check the voice quality of TTS systems:** The quality of all the languages can be checked and voice can also be saved from <http://www.iitm.ac.in/donlab/hts/>, visit to experience the TTS voice quality.

Reference :

http://tdil-dc.in/index.php?option=com_vertical&parentid=85&lang=en

Question 3: What are important components in Text to speech system And How this system is useful to build online Music Tool as part of e-learning.

Ans:

Important Components in Text To Speech System:

A text-to-speech system (or "engine") is composed of two parts: a front-end and a back-end. The front-end has two major tasks. First, it converts raw text containing symbols like numbers and abbreviations into the equivalent of written-out words. This process is often called *text-normalization pre-processing*, or *tokenization*.

The front-end then assigns phonetic transcriptions to each word, and divides and marks the text into prosodic units, like phrases, clauses, and sentences. The process of assigning phonetic transcriptions to words is called *text-to-phoneme* or *grapheme-to-phoneme* conversion. Phonetic transcriptions and prosody information together make up the symbolic linguistic representation that is output by the front-end.

The back-end—often referred to as the *synthesizer*—then converts the symbolic linguistic representation into sound.

In certain systems, this part includes the computation of the *target prosody* (pitch contour, phoneme durations), which is then imposed on the output speech.

Text-to-Speech (TTS) Tools for eLearning Applications

There are a number of possibilities available today for using TTS for eLearning; they fall into two categories or approaches:

1. **TTS Stand-Alone.** A general approach in which developers use any standard authoring tool such as Articulate or Lectora and use stand-alone TTS on-demand services/products to create audio files that are then linked or embedded in the presentation.
2. **TTS Integrated.** Products/services that have TTS voices bundled and integrated with an authoring solution, including Adobe Captivate and Tuval Software Industries' Speech-Over Professional.

TTS Stand-Alone Web Services

TTS stand-alone products can be used by eLearning developers irrespective of the authoring tool they are used. Several of the voice vendors offer on-demand TTS voice web services which accept text and produce sound files. Here are a few of the top web services for TTS:

These web services have the advantages:

- Choose any voice among a set of vendors voices
- Set pitch, speed volume of voice for the entire file
- Select type of sound file output (wav, mp3, etc)
- Preview function
- Pronunciation dictionary
- Pay as you go

Particular desktop stand-alone product to illustrate the eLearning production workflow:

1. Acapela Virtual Speaker – a Desktop Stand-Alone TTS Product
2. Personal TTS Readers Not Licensed for eLearning

References :

1. https://en.wikipedia.org/wiki/Speech_synthesis#:~:text=A%20text-to-speech%20system,equivalent%20of%20written-out%20words.
2. <http://elearningtech.blogspot.com/2010/07/using-text-to-speech-in-elearning.html>

Question 4: What are the list of Text to Speech systems based on their various methods, Techniques.

Different Systems Based On Text To Speech System :

Android

Version 1.6 of Android added support for speech synthesis (TTS).

Internet

Currently, there are a number of applications, plugins and gadgets that can read messages directly from an e-mail client and web pages from a web browser or Google Toolbar. Some specialized software can narrate RSS-feeds. On one hand, online RSS-narrators simplify information delivery by allowing users to listen to their favourite news sources and to convert them to podcasts. On the other hand, on-line RSS-readers are available on almost any PC connected to the Internet. Users can download generated audio files to portable devices, e.g. with a help of podcast receiver, and listen to them while walking, jogging or commuting to work.

Open source

Some open-source software systems are available, such as:

- Festival Speech Synthesis System which uses diphone-based synthesis, as well as more modern and better-sounding techniques.
- eSpeak which supports a broad range of languages.
- gnuSpeech which uses articulatory synthesis from the Free Software Foundation.

Others

- Following the commercial failure of the hardware-based Intellivoice, gaming developers sparingly used software synthesis in later games. Earlier systems from

Atari, such as the Atari 5200 (Baseball) and the Atari 2600 (Quadrun and Open Sesame), also had games utilizing software synthesis.

- Some e-book readers, such as the Amazon Kindle, Samsung E6, PocketBook eReader Pro, enTourage eDGe, and the Bebook Neo.
- The BBC Micro incorporated the Texas Instruments TMS5220 speech synthesis chip,
- Some models of Texas Instruments home computers produced in 1979 and 1981 (Texas Instruments TI-99/4 and TI-99/4A) were capable of text-to-phoneme synthesis or reciting complete words and phrases (text-to-dictionary), using a very popular Speech Synthesizer peripheral. TI used a proprietary codec to embed complete spoken phrases into applications, primarily video games.
- IBM's OS/2 Warp 4 included VoiceType, a precursor to IBM ViaVoice.
- GPS Navigation units produced by Garmin, Magellan, TomTom and others use speech synthesis for automobile navigation.
- Yamaha produced a music synthesizer in 1999, the Yamaha FS1R which included a Formant synthesis capability. Sequences of up to 512 individual vowel and consonant formants could be stored and replayed, allowing short vocal phrases to be synthesized.

Digital sound-alikes

With the 2016 introduction of Adobe Voco audio editing and generating software prototype slated to be part of the Adobe Creative Suite and the similarly enabled DeepMind WaveNet, a deep neural network based audio synthesis software from Google speech synthesis is verging on being completely indistinguishable from a real human's voice.

Adobe Voco takes approximately 20 minutes of the desired target's speech and after that it can generate sound-alike voice with even phonemes that were not present in the training material. The software poses ethical concerns as it allows to steal other people's voices and manipulate them to say anything desired.

E-Learning Tools:

1. Acapela Virtual Speaker – a Desktop Stand-Alone TTS Product
2. Personal TTS Readers Not Licensed for eLearning

References :

1. https://en.wikipedia.org/wiki/Speech_synthesis
2. <http://elearningtech.blogspot.com/2010/07/using-text-to-speech-in-elearning.html>

Question 5: What is a lexicon? How it is relevant term in Context of Text to Speech and Speech to Text systems?

Lexicon :

A **lexicon**, **word-hoard**, **wordbook**, or **word-stock** is the vocabulary of a person, language, or branch of knowledge (such as nautical or medical). In linguistics, a lexicon is a language's inventory of lexemes. The word *lexicon* derives from the Greek λεξικόν (*lexikon*), neuter of λεξικός (*lexikos*) meaning 'of or for words'.

Linguistic theories generally regard human languages as consisting of two parts: a lexicon, essentially a catalogue of a language's words (its wordstock); and a grammar, a system of rules which allow for the combination of those words into meaningful sentences. The lexicon is also thought to include bound morphemes, which cannot stand alone as words (such as most affixes). In some analyses, compound words and certain classes of idiomatic expressions and other collocations are also considered to be part of the lexicon. Dictionaries represent attempts at listing, in alphabetical order, the lexicon of a given language; usually, however, bound morphemes are not included.

Lexicon Relevant In Terms Of Text To Speech :

1. Multilingual lexicon dictionary for looking up words during translation while the software program allows effective and efficient interaction of components; and software programs to understand a process those rules.
2. Rule Based Model :
Transfer: It uses lexicons and structural analysis into every SL input text after which it's converted to intermediate representation.
3. Hybrid machine translation is a combination of RMBT (Rule Based Machine Translation) and SMT (Statistical machine translation) method, and it makes the use of the advantages of both these methods. Statistical data is hence, put to use in generation of lexicon and syntax.

References :

1. <https://en.wikipedia.org/wiki/Lexicon>
2. <http://www.igntu.ac.in/eContent/IGNTU-eContent-815947141046-MA-Linguistics-4-HarjitSingh-ComputationalLinguistics-5.pdf>

Question 6: What is an OCR? How it is related to the Text to speech system. List a good Journal and Conference papers for recent years published i.e last one decade from the today date. i.e 3rd November 2020 batch to 2010 January.

Answer:

OCR : Optical character recognition or optical character reader (OCR) is the electronic or mechanical conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image (for example: from a television broadcast).

OCR Related To Text To Speech System :

In text to speech module text recognised by OCR system will be the inputs of speech synthesis system which is to be converted into speech in .wav file format and creates a wave file named output wav, which can be listen by using wave file player.

Two steps are involved in text to speech synthesis

- i) Text to speech conversion
- ii) Play speech in .wav file format

1. **Text to speech conversion: In the text speech conversion input text is converted speech (in LabVIEW) by using automation open, invoke node and property node.**
2. **Speech synthesis :A wave file output.wav is created containing text converted into speech which can listen using wave file player.The waveform will vary according to the different text from OCR output in the text box and can be listened on the speaker.**

Research And Conference Papers :

1. <https://www.semanticscholar.org/paper/REAL-TIME-IMPLEMENTATION-OF-IMAGE-RECOGNITION-AND-Kalaivani-Praveena/99c990ff841da4923771c5d2b6692c3a4656d447?p2df>
2. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.682.5362&rep=rep1&type=pdf>
3. https://link.springer.com/chapter/10.1007/978-3-642-31522-0_71
4. <https://ieeexplore.ieee.org/abstract/document/7919526/>
5. <https://ieeexplore.ieee.org/abstract/document/8284628>
6. <https://patents.google.com/patent/US7805307B2/en>

Question 7 : List out Indian language specific Text to speech system that are developed by Industry and Academic research centers that work in the domain of Speech and signal processing with a tabular form. Also list out the which of

Indian language Speech to Text systems are working well for any societal applications.

Answer:

Indian Text to Speech System:

1. The Indian Government has **launched** Text-To- Speech system (TTS) computer software in six Indian languages namely: Hindi, Marathi, Bangla, Telugu, Tamil & Malayalam, and web based Optical Character Recognition systems (OCR) for Hindi (Devanagri) & Punjabi (Gurumukhi). Text-to-Speech converts computerized text into human voice, allowing people with visual impairments to listen to written works on a computer or a mobile device. TTS and OCRA have been developed under a consortium project by Technology Development for Indian Language (TDIL), a program under the Department of Information Technology.
2. The Text to Speech system integrated with a screen reader would also enable visually challenged users to interpret and perform computer operations with an audio interface. The Optical Character Recognition systems convert hard copies to electronic form so that printed text can be accessed on a computer. OCR for Indian scripts can make this literature web accessible. OCR can also be used to generate Braille version of printed books. OCR's can also enable text-to-speech systems to deliver Indian language books through the medium of audio.
3. The TTS has also been integrated with Non Visual Desktop Access (NVDA) screen reader and Optical Character Recognition System (OCRA).
4. The Government of India had also launched the Indian Language Technology Proliferation and Deployment Centre (ILTP-DC) portal, under the Technology Development in Indian Language (TDIL) program, offering a single window access system to various linguistic resources and software tools for Indian academic researchers.

Languages That Are Working Well In TTS System:

1.Hindi

2.Marathi

3.Bangla

4.Telugu

5.Tamil

6.Malayalam,

and web based Optical Character Recognition systems (OCR) for

- 1.Hindi (Devanagri)
- 2.Punjabi (Gurumukhi).

References :

1. [https://www.medianama.com/2011/09/223-indian-governement-launches-text-to-speech-indic-ocr-systems/#:~:text=The%20Indian%20Government%20has%20launched,\)%20%26%20Punjabi%20\(Gurumukhi\).](https://www.medianama.com/2011/09/223-indian-governement-launches-text-to-speech-indic-ocr-systems/#:~:text=The%20Indian%20Government%20has%20launched,)%20%26%20Punjabi%20(Gurumukhi).)

Question 8: What kind of programming language skills are necessary to develop the TTS system? What kind of subject relevant knowledge required to develop your own language specific Text to Speech system with your voice as good as natural speech. Give good references from any reputed Journals and Good International Conference research papers, demos.

Ans:

Programming Skills :

A good **language** for this task is Python. You **can** use a library called CMU Sphinx, and when you integrate it with python, it is relatively simple. This library works with numerous **languages** some of which being Java, Ruby, and Scala.

Language Knowledge:

Strongly recommend reading up on the basics of signal processing first. This will enable you to understand how the Python code works and make you a better NLP and deep learning professional!

Good Research Papers and Journals:

1. https://d1wqtxts1xzle7.cloudfront.net/30521305/inproceedings.pdf.cd05a294-e2bf-405f-a042-19d9d8e0a388.pdf?1359882001=&response-content-disposition=inline%3B+filename%3DText_Processing_for_Text_to_Speech_Syste.pdf&Expires=1604668246&Signature=XhNTU86A9PxraD-0ILDKMb3sgF6NbMAQUE8dHDHEIEondaTNLd1qJUrMzbScx9~MYUJTd9Xce5cyWTLHlf~idIwmjvV4SwWZ5JAnfnXrmRNcbQInfxdprfwpdq43CVBkRcSj7s7vb9jZDMB

[FVkuF-rvZsNqe09zmUbPQHH15HLi7Mr7iaOdfP~IXfU~onx1uiZcJ4PIMUZah5J5g8KRimhy9FupBGSPbNJ5FphndSX7inZA9bfNWobR~EcPM0akbuEkROTPCFk5x2YE1LJL4LgN~1HrvlyWH99svu3hh3vAynz1Gx-PreLI8gWrmI6CfW3A5cFhlE9QJrc6RrxkTng &Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA](#)

2. <https://patents.google.com/patent/US6535852B2/en>
3. https://link.springer.com/chapter/10.1007/11846406_55

References :

1. https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=Building+Text+To+Speech+System&btnG=
2. <https://www.quora.com/How-can-I-learn-programming-for-voice-recognition#:~:text=Simple%20Answer%3A%20A%20good%20language,Java%2C%20Ruby%2C%20and%20Scala.>
3. <https://www.analyticsvidhya.com/blog/2019/07/learn-build-first-speech-to-text-model-python/>

Question 9: List out in a tabular form based on various technology based implementation Text to speech System. Give minimum 10 different tools with their detailed description about its functioning and process.

Ans:

Different Tools About Text To Speech :

1. **Ivona** : A member of the Amazon group of companies, Ivona is one of the best text to speech software tools in the market.
 - **47 Human-Like Voices**
Ivona covers 24 languages with both male and female voices for almost all of them.
 - **Quality Performance Regardless Of Device**
Ivona partners with ReadSpeaker bringing optimized performance on almost all devices, as well as the ability to integrate Ivona into other applications.
2. **NaturalReader**

Another great text to speech software with Optical Character Recognition for both Windows and Mac users. NaturalReader also offers the ability to change speed of speech.

- **Free Download With Great Compatibility**
Free version available and compatible with PDF, Word files and Webpages.
- **Vast Range Of Voices**
US English, British English, French, German, Italian, Spanish, Swedish, Arabic, Chinese and Japanese to name a few.
- **Optical Character Recognition (OCR)**
OCR enables you to scan your text and then convert it into digital text which you can listen to in audio form or save in your computer.
- **Great Aid For People With Learning Disabilities**
From dyslexics to visually impaired online learners the “see it highlighted as you listen to it” function is a one key fits all solution. Real-time spelling feedback further strengthens spelling weaknesses. Online learners with [learning disabilities](#) can benefit greatly from such features.

3. [Zabaware Text To Speech Reader](#)

Great text to speech software with a speech synthesizer that reads many types of documents.

- **Various Compatible Text Types**
Documents, emails, websites, clipboard contents as well as Windows dialog boxes.
- **Free Download Available**
Free version comes with high-quality non-human voices.
- **Great Tool For Learning Disabilities**
Reading disabilities, as well as various concentration problems, are tackled by Zabaware’s features.

4. [iSpeech](#): **iSpeech is a text to speech software with the ability to convert text to numerous audio formats not limited to only one device. Blackberry, iPhone and Android applications are available.**

- **Multiple Formats**
Wav, MP3, ogg, wma, aiff, alaw, ulaw, vox, MP4 to name a few.

- **Give Voice To Any Text**
Don't limit yourself to PDFs or copy and pasted text. iSpeech's open-source text to speech function allows you to voice-enable even text from chat applications.
- **Multiple Devices**
Nobody owns just one mobile device nowadays. iSpeech has you covered with its responsive design.

5. [Acapela Group Virtual Speaker](#)

One of the best [text to speech software](#) tools the market has to offer, particularly useful for eLearning purposes, with many compatible formats, languages and voice properties.

- **Immense Variety Of Voices**
More than 70 voices and 30 languages with additional voice mood range as well as voice frequency to indicate sadness, happiness, whispering or screaming.
- **Several Audio Output Formats**
You name it, Virtual speaker has it. 8 kHz, 11 kHz, 16 kHz, 22 kHz, 44 kHz along with many more.
- **Speech Control**
Modify settings, voice or language as you are listening to a text with this great feature.

6. [TextSpeech Pro](#)

Download-only text to speech software that reads the most popular document types such as PDFs, MS Word files and HTML. It also provides support to impaired users.

- **Great Selection Of Natural Voices**
The paid version includes AT&T's natural voices with the ability to change voice speed, volume and quality.
- **Various Integration Features**
TextSpeech Pro gives you the option to have your emails read out loud, as well as have any web page read to you.

7. [AudioBookMaker](#)

If you're on the lookout for free and reliable text to speech software, then look no further. AudioBookMaker is probably the best free text to speech software.

- **Multi-Lingual Interface**
English, Russian and simplified Chinese included.
- **Customizable Speech Parameters**
Change speed, pitch and volume as you see fit.
- **Highlighted Spoken Text**
Significantly important for online learners with learning disabilities.
- **Customized Settings**
Not only can you adjust a variety of document settings, but you can also have those settings restored every time you re-open that document.

8. [TextAloud 3](#)

Developed by NextUp, TextAloud 3 is one of the most professional text to speech software tools, featuring 29 languages.

- **Massive Range Of Voices**
TextAloud has partnered with AT&T Natural Voices, Acapela Group, Ivona, and Nuance Vocalizer bringing you many voice options and different accents to choose from.
- **Free Trial Version**
You don't have to pay straight ahead. You can download the free trial and then decide if you want to move on with a premium subscription.
- **Tons Of Features**
Create messages for your answering machine, have your text read in 29 languages, or use TextAloud 3 as help for people with sight problems.

9. [Read The Words](#)

Online text to speech software with various language options and easy-to-use interface with free version available.

- **3 Language Options**
Have your text translated and read to you in English, French and Spanish.
- **RSS Audio Podcasts**
Articles from an RSS feed are converted into speech recordings, while the links are kept in the feed. There is a 3-day trial available so you can test out this great feature.
- **Audio File Creation**
Free users can create audio files of up to 30 seconds.
- **15 Voices**
Free version offers 15 different voices to choose from.

10. **Voice Reader 15** : Linguattec has produced this excellent text to speech software tool with numerous functional features.

- **45 Languages**
English, German, Italian and Spanish among others, depending on the membership you choose.
- **Translator**
Also available in many premium memberships to suit either individual or corporate needs.
- **Mobile Applications**
Plethora of apps such as the Voice Translator for interpretation purposes, text to speech mobile applications, VoiceReader and VoiceReader Web for both PDF documents and webpages. Also, the LinguaDict integrated voice output German dictionary covers all your lexical needs.

List Different Technologies Based On Tabular Form Based On Text To Speech:

- Natural Reader
- Snap & Read
- Word Q
- Kurzweil 3000
- Text Help Read and Write

References :

1. <https://www.readingrockets.org/article/text-speech-tts>

2. <https://elearningindustry.com/top-10-text-to-speech-tts-software-elearning>

Question 10 : What is the different fundamental and derived component that are possible to extract from the Praat tool during feature extraction method of Music data from your practical exercise submitted as class assignment. List them and describe in detail its process and functions.

Ans:

1. Pulse :

To add:

click at the desired time location, and choose **Add pulse at cursor** or type **Command-p**.

To remove:

make a [time selection](#), and choose **Remove pulse(s)** or type **Option-Command-p**. If there is no selection, the pulse nearest to the cursor is removed.

Pulse (this allow you to set pulses (necessary for e.g., pitch analysis) and to extract specific information on voice parameters like jitter and shimmer; pulses are indicated in the top panel with vertical dark blue solid lines)

The 'pulses' menu also lists many different jitter and shimmer parameters that are available in PRAAT. A detailed description of these measures can be found in the 'Voice' section of the main manual (where these measures are compared to the Multidimensional voice program (MDVP) parameters of the Kay company).

2. Formant:

Formant Pro is a convenient tool built for large-scale, systematic experimental studies of formant movements. It allows users to obtain continuous formant and formant velocity trajectories from multiple sound files, take various measurements, and save them in formats ready for graphical and statistical analysis. It also generates averaged trajectories and measurements across repetitions and speakers.

There Are Many Forms In There :

1.f1

2.f2

3.f3

3. Intensity :

Measuring intensity in Praat is relatively straightforward, albeit with a major disclaimer. To get the overall intensity of a sound, select the desired sound and run Objects → Query → Get Intensity (dB). To get the intensity at a specific point in the sound, open it in an editor window, Editor → Intensity → Show Intensity, and then use the various commands available in the Editor → Intensity menu to get whatever information you desire. By default, Praat's display of the intensity of a word is smoothed to avoid showing individual pulses in the amplitude lines, both in the editor window and in amplitude objects (when drawn or viewed). This smoothing is based on the minimum F0 of the sound. If you want to see something closer the amplitude envelope of the sound in Praat (where pulses show up individually as amplitude peaks), or if you want the amplitude curve to be smoother than it normally would be, you must simply adjust the minimum pitch expected by Praat. This can be done in Editor → Pitch → Pitch Settings..., as described more fully in Section 6.4.3. Similar smoothing/desmothing can be accomplished when creating Intensity objects by altering the minimum pitch value in the Objects → To Intensity... dialog box. This decrease in amplitude smoothing is particularly useful for measuring or counting quick, amplitude-based phenomena like taps and flaps. That said, in most recordings made for phonetic research, absolute intensity measures as given by Praat are largely meaningless. To accurately measure the absolute intensity of a speaker's voice, a sound-attenuated booth with a calibrated sound level meter or calibrated microphone with specialized software should be used. Relative intensity (say, between two segments or words) can be measured with an uncalibrated microphone, but is only accurate if the recording is made in a consistently quiet area, and the speaker stayed in the same general position relative to the microphone throughout the recording(s) (and wouldn't have changed much during the time between the two points of comparison). This issue is discussed in depth in Praat's user manual.

4. Pitch:

1. Measuring F0 from a single cycle :

The surest way to get an accurate F0 for a single cycle is to open the file in the Editor window, then: 1. Zoom in to the point where you can see individual cycles in the sound file 2. Select one complete cycle, as accurately as possible, thus, giving Praat the period in seconds (t) 3. Praat will calculate the frequency of the sound in Hertz in the top bar, giving it in the format (____/ s). Use the zoom sel button to zoom in if you can't see the frequency readout.4 6.4.2 Viewing Pitch via a narrowband spectrogram The most reliable way of getting a sense of the pitch through the course

of the word in Praat is by examining a narrowband spectrogram with a reduced visible range (0 - 400 Hz for speech).

This can be done by editing the spectrogram settings as described in Section 6.2. The contours of the harmonics will accurately represent the pitch contours of the voice during the word, and doing this will give you a sense of the contour before using the Praat pitch tracker for more precise measurement. 6.4.3 Using Praat's Pitch Tracking Praat does have the ability to provide a pitch track in the editor window. To enable the pitch track in the Editor window: Editor → Pitch → Show Pitch At this point, a blue line will be placed on top of the spectrogram representing the pitch, where Praat can find it. Once the pitch track is placed, you can use the cursor to check the pitch at any given point in the word. Just place the cursor and look for the middle blue number on the right side of the window. You can also place your cursor at a given point in the file and Editor → Pitch → Get Pitch. Running Editor → Pitch → Get Pitch when a chunk of the sound is collected will return the average pitch during that selection.

2. Improving Pitch tracking by changing the Pitch Settings

Then Change These Settings:

1. Pitch Range (Hz)
2. Unit
3. Method
4. Silence threshold
5. Voicing threshold
6. Octave Jump cost
7. Voiced / unvoiced cost

3. Scripting: Creating a Pitch Object

4. Getting Maximum, Minimum, and Average pitch for a section of speech

5. Spectrum :

Sometimes, you need specific details about the frequencies and individual harmonics in a sound at a given moment in time, and examining a narrowband spectrogram alone does not provide sufficient information. In these cases, you'll need to take a spectral slice for analysis. Spectral slices (also referred to as FFTs or spectra) are the result of a fast fourier transform done on a very small portion of the sound, providing you with very specific information about the frequencies present in the sound and their relative amplitudes. Spectral slices are useful for a variety of measures of F0, nasality, creak, breathiness, and spectral tilt, and are a crucial part of many measurement workflows.

1. Taking a spectral slice To take a spectral slice, you'll need to do the following:

1. Editor → Spectrum → Spectrogram Settings
2. Set Window Length to “0.025” (effectively producing a narrow-band spectrogram)
3. Editor → Spectrum → Advanced Spectrogram Settings
4. Set Window Shape to “hamming”
5. Select the point at which you’d like to see the slice taken
6. Editor → Spectrum → View Spectral Slice

This will create a new Spectrum object, and pull up a window like that in figure 5, showing amplitude on the Y axis, and frequency (from 0 up to the Nyquist frequency) on the X axis. You can zoom in and out using the buttons in the bottom left corner, as you wish. If you’ve selected a portion of the sound (rather than a single timepoint) when you use Editor → Spectrum → View Spectral Slice, Praat will create a spectrum representing the average characteristics across the entire selection.

