

PROSODY

Prof. Dr. Eşref ADALI

Chapter – III

E-mail : adali@itu.edu.tr

www.adali.net or www.xn--adal-oza.net

What is Prosody?

- The use of alphabet has made a significant contribution to the transfer of spoken words to writing, but we cannot say that it is sufficient. Because there is no feature of expressing emotion in a written text.
- We can convey our feelings to the other person by changing the amplitude of our voice, emphasizing or prolonging certain words during speech. In addition, the features and tone of our voice add character to speech.
- Prosody studies how such features of speech can be translated into writing.

From a grammatical point of view, it is about the rhythm, emphasis and intonation of speech. It indicates the following qualities of the speaker:

- emotional state,
- What you say is a discourse, a question or an order,
- The way of saying it is sarcastic, sarcastic, emphatic, contradictory and focused.

In terms of acoustics, it deals with

- the length of the sounds,
- the pitch of the voice,
- the pitch of the speech, and
- the formant frequencies.

Emphasis

When a word or sound is pronounced more distinctly, stronger and louder than other words or sounds while speaking is called emphasis. As it can be understood from this definition, stress can be pronounced with a phoneme in a word, as well as a word in a sentence.

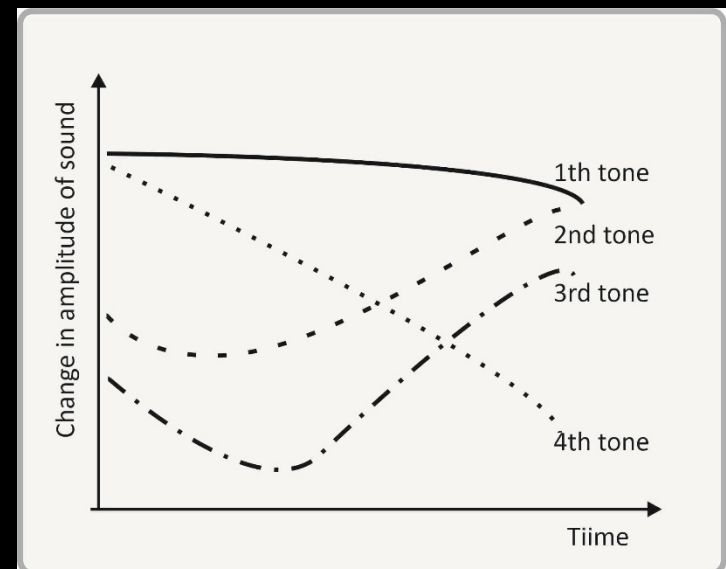
Intonation

- We change the amplitude of our voice while speaking some words and sounds. In other words, we play with the amplitude of our voice. Such changes affect the meaning of the word.
- In some of the world's languages, the influence of intonation is so great that they are referred to as **tonal languages**. The term **pitch** is also used instead of the term tone.
- Chinese and Swedish are counted among the tonal languages. In Mandarin, the way a word is pronounced, changes the meaning of the word.

Level of sound	Direction of sound level
┐ Veri hight	↗ Increasing
┐ Hight	↘ Decreasing
┐ Middle	↗ Hight and increasing
┐ Low	↗ Low than incresing
┐ Very low	↗ Increasing than decreasing
↓ Decreasing	↗ Genrally increasing
↑ Increasing	↘ Generally decreasing

IPA Symbols for intonation

- When said in the 1st tone, it means: **mother**
- When pronounced in the 2nd tone, it means: **hemp**
- When said in the 3rd tone, it means: **horse**
- When said in the 4th tone, it means: **to scold**



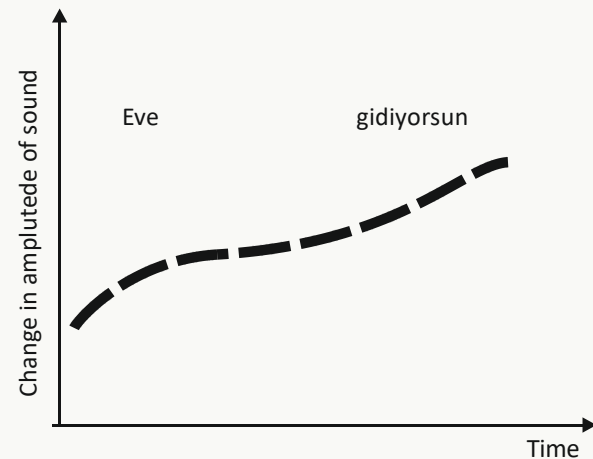
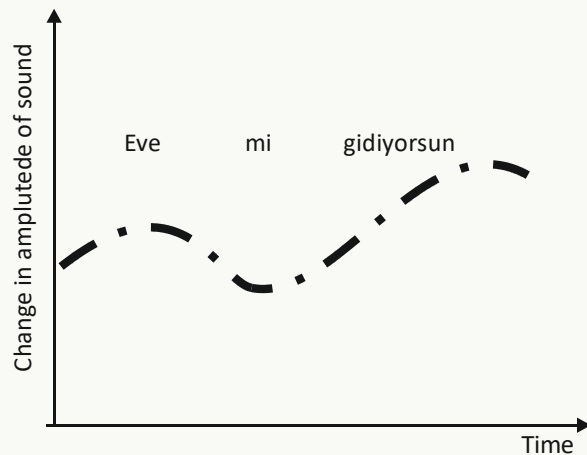
Mandarin

馬

mǎ (horse)

Jitter and Shimmer

- When we speak, we do not always keep the level of our voice at the same level, we lower or raise it. The pattern of change in the speaker's voice during speech is called the melody of speech. The tune is important in terms of understanding the speech and making an impact on the listeners.



Duration

- Some solutions have been searched to be able to read a vowel in a word longer than the other vowel. However, the solutions found are language specific and cannot be said to be exact solutions.
- Two solutions have been found in Turkish alphabet for this issue: The first solution is provided with "ğ", "ğ" provides long reading of a vowel sound before it.
- However, when it is used as a merging letter in the addition of possessive suffixes, the letter **k** becomes a soft equivalent. It provides long reading of a and e sounds in dağ, yağ samples. It provides softening of the consonant in the examples of **yemek** (**yemek-i**) and **uçak** (**uçak-ı**).

The second solution is to put a caret on the vowel. In some cases, the caret lengthens the vowel, in some cases it thins the consonant, and in some cases it undertakes two functions together. The caret mark placed on the letters "a" in the word still ensures the lengthening of the first sound and the thinning of the "l" sound. On the other hand, the hat in the word profit allows the letter "k" to be read delicately.

In some alphabets, there are structures for bodily purposes. For example, there are three letters in Arabic that make the letters that come before them read long. These are the letters **elif**, **vav** and **ye**, they are called extension letters because of these features.

دَامَ

dâme

بِه

bîhi

دُون

dûne

رَبِّه

Rabbuhû

دِينَ

dînin

Stop

- The speaker is silent for a while after the word to be emphasized during the oral presentation.
- Thus, he emphasizes the word he cares about to the audience. (-) sign will be used to indicate the stop point.

- We are opening - the world's largest dam today.
- He can give as much as he wants - to his brother.
- In a bikini - no swimming.

Emotion

The tone of voice of the speaker is a reflection of the emotional state he is in. Possible situations could be:

- state of calm
- state of being sad
- fear state
- state of being angry

In these four cases, there is a change in the tone of the voice and the amplitude of the voice.
E.g:

- He raises his voice because he is angry.
- He raises his voice and changes his pitch because he is angry.
- He increases his speaking rate when angry.

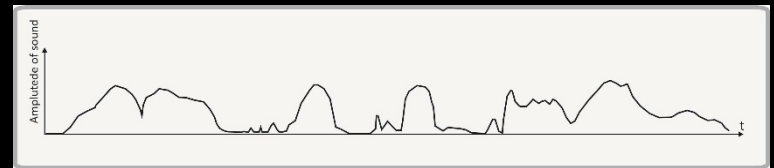
Assuming the speaker has four emotional states. A sad person's speech is smooth and low in volume. The ups and downs are high in the speech of a frightened person. The loudness of the voice is also quite high. It is seen that the loudness of a calm person is at a medium level. Also, there are not many ups and downs in his speech. An angry person will have a loud voice and have strong ups and downs in the voice.

From Acoustic View of Prosody

Studies on the emotional state of the speaker and acoustic measurements are available in different sources. A study by Williams and Stevens in 1972 was carried out on three people. Three people were told the same words in four emotional states and their voices were recorded. Then, these records were evaluated from a technical point of view. The test results are as follows:



electrical signals of the sentence



change in the amplitude of the sentence

- The frequency of the sound produced by each speaker has a middle value. This is called the fundamental frequency f_0 .
- The value of f_0 may vary from speaker to speaker.
- The frequency of the speech changes according to the emotional state of the speaker. Frequency change is less during calm and sad speech, and more in anger and fear.
- The basic frequency of speech also changes according to the emotional state of the speaker. The fundamental frequency is lowest in the sad state, slightly higher in the calm state, slightly higher in the fear state, and highest in the angry state.

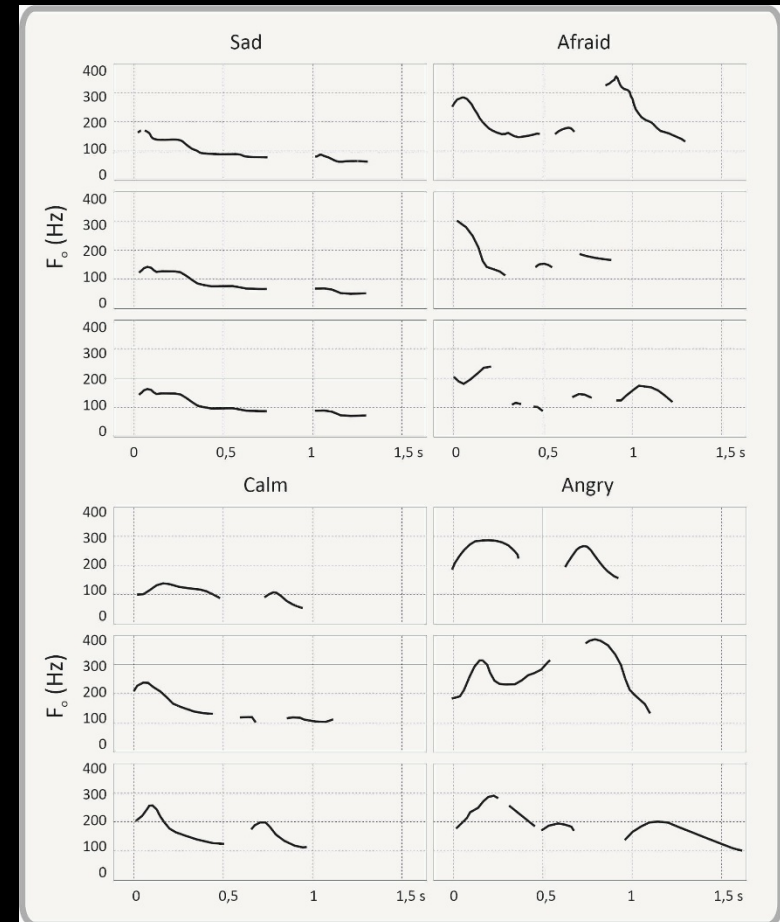
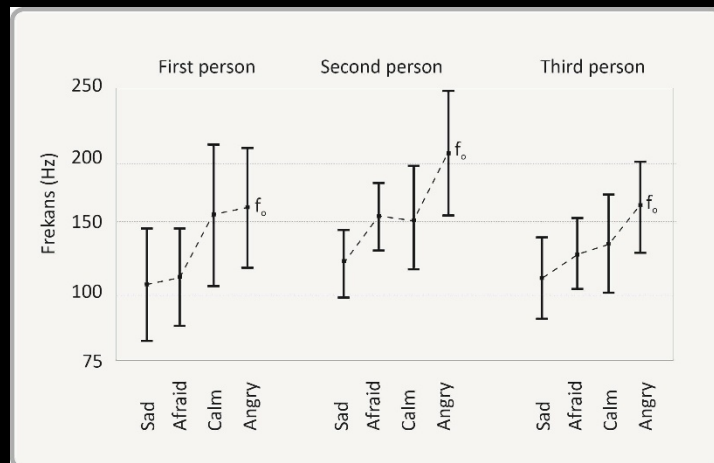
Result of Test

1- When the three speakers are calm, the speech melodies can be considered close to each other. The difference between the fundamental frequencies of their voices may be due to the mouth structure and gender of the subjects.

2- In the case of sadness, the change in the amplitude of the fundamental frequency of the voice remains relatively small.

3- In the case of fear, the change in the amplitude of the fundamental frequency of the sound is quite large compared to the calm state. In addition, the threshold value of the fundamental frequency is slightly higher compared to the calm state.

4- In the case of anger, the amplitude of the fundamental frequency of the sound becomes very high. Likewise, the fundamental frequency threshold of the sound is also considerably higher.



References

- [1] N. Chomsky, M. Halle, The Sound Pattern of English, Harper&Row, 1968
- [2] C. E. Williams, Emotion and Speech: Some Acoustical Correlations, The Journal of the Acoustical Society of America, Vol:52, No:4, 1972
- [3] P. Pattnaik, S. Dash, A Study on Prosody Analysis, International Journal Of Computational Engineering Research (ijceronline.com) Vol. 2 Issue. 5, Sept. 2012
- [4] J. Clark, C. Yellop, J. Fletcher, An Introduction to Phonetics and Phonology, Blackwell, 2007
- [5] G. Güneş, Limits of Prosody in Turkish, 2015, <https://www.researchgate.net/publication/281235927>
- [6] G. Güneş, On The Role of Prosodic Constituency in Turkish, <https://www.researchgate.net/publication/281235850>
- [7] S. Nooteboom, The Prosody of Speech: Melody and Rhythm, Research Ins. for Language and Speech, Utrecht Univ Holland
- [8] T. Johnstone, The Effect of Emotion on Voice Production and Speech Acoustic, Doktora tezi, University of Western Australia & University of Geneva, 2001
- [9] Y. Xu Speech Prosody — Theories, Models and Analysis, Cambridge Scholars Publishing, pp 171-204, 2012
- [10] M. O. Külekçi, K. Oflazer, An Infrastructure for Turkish Prosody Generation in Text-to-Speech Synthesis, <https://pdfs.semanticscholar.org/.../26ac334a002e3f96d951>.