AN OVERVIEW OF INTONATION CONTOURS AND SPEAKING RATE IN THE SPEECH OF AN INDIVIDUAL WITH FLACCID DYSARTHRIA

Abebayehu Messele

ABSTRACT

It is an established fact that language is the chief distinctive feature of humans. Unfortunately, there are a number of individuals missing usage of language and/or speech partially or entirely, from the beginning or lately, momentarily or for a long time, usually for defined reasons or for reasons which cannot yet be explained. This paper attempts to investigate the extent of abnormality in intonation patterns and speaking rate of the speech of an individual with Flaccid Dysarthria. The subject was diagnosed as a Primary Lateral Sclerosis patient, which is a progressive degenerative motorneuron disease, that is, nerve cells in the body gradually die off. It affects only some of the nerve cells in the body—those that control voluntary movement of muscles. The main objective of this study is to acoustically analyze the patient’s vowels, consonants, intonation contours and duration.

Acoustic methods were employed to examine the intonation contours and speaking rate of the patient’s speech. More specifically, the data were from solicitation and spontaneous utterances of the subject. They were digitally recorded, sampled and quantized, then fed into a speech analyzer software called Praat. The interpretations of the data were done on the basis of the facts revealed by the software. Accordingly, the patient’s intonation contours and speaking rate were found to be deviant.

I. Introduction

Consonants and vowels are referred to as segmental because, in principle, any utterance can be segmented into a finite number of non-overlapping consonants and vowels. Segments follow each other in time. Yet, speech involves more than stringing together these individual segments in a sequence. Suprasegmental features (prosodies) are other aspects of speech laid on top of a segment or group of segments.

---

1 Dysarthria is a motor speech disorder which largely happens due to breakdowns in movement control of one or more muscle groups that compose the speech mechanism. It could be caused, among other things by Parkinson's disease, stroke, brain injury, tumors, cerebral palsy, Lou Gehrig's disease, Huntington's disease.

2 I am grateful to the subject of this work by the name Yohannes Abreha, who had been very helpful during data collection and medical examination. I am also very much indebted to Dr. Moges Yigezu for his supportive suggestions.
We define segments in terms of place and manner of articulation, states of the glottis, and air streams mechanism, but we can not do the same for suprasegmentals because they are supplementary elements that change the voice quality of utterances. The idea of prosody (suprasegmental) basically includes duration, stress, tone, pitch accent and intonation. The objective of the present study is not to treat all of these notions. Rather focusing on one of them, that is, intonation with reference to the patient's speech.

1. Intonation: Types and Uses

By intonation is meant the way physical or prosodic parameter of fundamental frequency is perceived as regular pitch patterns across a sequence of speech units (Botinis, 1998). In other words, there is a constant change in the pitch of the voice when we utter sentences. The difference between speaking and singing is that in singing we hold a given note for a noticeable length of time and then jump to the pitch of the next note. But in speaking, there are no steadystate pitches (Ladefoged, 2001:99). The intonation of a sentence is, therefore, the pattern of pitch changes.

Different literature suggest that there are, in general, about four types of intonation pattern so far identified. These are: Rising intonation, that is, the pitch of the voice increases over time; falling intonation, that is, the pitch decreases with time; dipping intonation, that falls and then rises; and peaking intonation that rises and then falls. Note, however, that these intonation types are not the only ones; there are many different contours of intonation in languages.

Intonation is often thought of as being directly linked to the speaker's emotions. No doubt that emotions are expressed through intonation levels and it is for this reason that intonation is often referred to as iconic and must be studied in relation to the entire gestural setting, particularly in relation to facial expressions and expressive body languages. Lexical and grammatical meanings, along with body languages and all accompanying physical expressions, are expressed through intonation as well. Grammatical meanings like focus, for example. In Amharic (and in English too), contrastive emphasis is marked using an intonational accent. Consider the following sentence, for instance, /läm isajə dəbbə ?ifälligallähu/ which means 'I want bread for my lunch' (as opposed to 'injera'—a traditional Ethiopian staple food). In producing this sentence the speaker exerts extra pitch and therefore intensity on 'bread' as marking emphasis.

2. Some Notes on Amharic Intonation System

In order to clearly see the problems of intonation patterns of the patient's speech it is very important to have some understanding of Amharic intonation system. Alemayehu (1987) has labeled Amharic as an intonational language based on the function of pitch. He has also made a detailed description of intonation contours of different types of declarative, imperative and
interrogative sentences. In this work, only some of the declarative and interrogative sentences of the patient in comparison with the normal trend will be examined.

With regard to declarative sentences, there are at least five different ways of supplying the declarative sentences with intonation (for further discussion, see Alemayehu, 1987:106). The pattern of the pitch differs depending on which constituent of a statement is given more prominence. As mentioned above, among the five types of declarative sentences presented by Alemayehu, this work only considers what he calls Declarative I. The contour of Declarative I sentence is a simple assertion statement. The pitch begins at a somewhat low level and rises on the penultimate word and gradually falls to the speaker's base line. And declarative II sentence is represented by emphasis that the speaker wishes to put on a constituent of a sentence and by the rise associated to the emphasis (ibid). It is believed that these types of declarative sentences are enough to see the deviation of the patient's speech.

The other type of sentences to be considered is interrogatives. Like Alemayehu (1987:169) claims, in Amharic, there are two different ways of making question sentences: Either by making use of question words like /mɨɨn/ 'what', /lämɨɨn/ 'why' and so on; or by assigning specific intonational features to otherwise structurally affirmative constructions resulting in the different Yes/No questions (ibid). Though there are different types of questions constructed by using question words and intonation pattern, in this work, we will only be looking at one sentence for each. This is because the main purpose of this work is just to show the fact that the patient's speech is deviant with respect to intonation and not to describe all of his intonation patterns in different types of sentences.

1. **Intonation Patterns of the Patient's Speech as Opposed to the Normal**

With the discussion of intonation types and their associated meanings let us now examine what the patient's intonation patterns look like. Four utterances were extracted from the subjects' speech for analysis. Two of them are declarative sentences, and the other two are interrogative sentences. Below are waveforms and Fo trajectories (fundamental frequency graph) of an assertive sentence (i.e., Declarative I).
Figure 1 Waveforms and pitch graphs of a sentence /tät'ärətt'ärku/ glossed as 'I doubted' produced by the patient (a) and by the normal person (b).

In the figure above, the pitch graph, below the waveform, designates the pitch contour of the sentence. In this figure, the patient produced the statement /tät'ärətt'ärku/ just to declare the fact that he was not certain of something. In other words, the sentence, according to Alemayehu (1987:107), falls under Declarative I. The normal pattern of intonation for Declarative I, as mentioned above, starts with a fairly low pitch and finally falls to the speaker's base line. This is exactly what is seen in (b); but in (a), the pitch starts to go downwards at the start of the last syllable, that is, /-rku/ and as a result it is not loud enough to be heard. In fact, /-r/ is in a better situation than /-ku/. In other words, towards the end, the pitch of the last syllable falls even below the patient's base line.

The fact that /-rku/ is not audible in the sentence being analyzed reveals one important problem about the patient's speech, that is, he is not able to produce sentences with their complete meanings. Because Amharic is one of the languages whose morphologies are very complex; and therefore there are elements, in these languages, that are suffixed and mark grammatical meanings. So, in the patient's speech, it is possible and even probable for these elements not to be heard, which, in turn, makes his speech less intelligible. In this particular case, for example, /-ku/ marks person and number, as mentioned in the footnote, and yet does not have

---

3 This statement seems just one word because all the elements that mark different grammatical meanings are fused together. The suffix /-ku/, for example, marks person (i.e., first person) and number (i.e., singular).
sufficient pitch that would enable it to be heard by normal human ear.

Technically when one sound has an intensity of 5 dB greater than another, it is approximately twice as loud; and a change in intensity of 1 dB is a little more than the just noticeable difference in loudness (Ladefoged, 2001:165). In the above sentence, the mean intensity is 81 dB, and the mean intensity of the last syllable is 75 dB. It appears, therefore, that the last syllable is more than twice quieter than the average intensity for the entire sentence.

Similarly, the patient’s /tt/, too, has less intensity than the average. This shows that his speech is inaudible even for noticeable sounds like ejectives. And, as will be discussed latter, geminated consonants have a relatively higher pitch (therefore higher intensity) than their ungeminated counterparts. Yet in his production of geminated consonants, too, we observe similar problem of inaudibility. So his sentence can be written as [tät'ärott'ärku].

The other case where this deviation is also revealed more is in another declarative sentence (Declarative I): /jämäslänɲɲɑl/ glossed as 'I think so.' Below are the waveforms and pitch graphs of the sentence as produced by the patient (a) and the normal (b).

![Figure 2](image-url)  

**Figure 2** Waveforms and pitch graphs of a sentence /jämäslänɲɲɑl/ glossed as 'I think so' produced by the patient (a) and by the ‘normal’ (b).

In the contours of the two productions of /jämäslänɲɲɑl/, there are at least two noticeable differences displayed in the pitch graphs.

---

4 Rising and falling intonations are marked by diagonal arrows—rising is marked by [↗] and falling by [↘].
Firstly, the patient started the sentence with a raised but immediately falling pitch, but the normal pattern starts with a relatively low but rising pitch. Secondly, in (a) the pitch is raised all of a sudden at the middle of the second syllable (i.e., /-mäs-/), and immediately goes down at the end of the syllable. In relation to the second one, one might think that it is a kind of signal processing error committed by Praat, i.e., the speech analyzing software. However, it has been checked that it is not an error. This was done by zooming into the part of the signal that sounds abnormal and checking the flow of the harmonic waves. From what has been discussed so far, we can tell that the contour for a Declarative I sentence is abnormal. Let us now consider what the situation looks like in the case of questions.

The above figure represents a question extracted from the patient’s narration. He was asking the researcher the question with an expectation of 'yes' or 'no' answer. As mentioned above, final falling pitch marks a ‘yes/no’ type of question in Amharic when the sentence is in a question form. The patient's intonation (i.e., (a)), seems to be normal and yet there is one deviation that we observe, that is, the Fo declines at the end of every intonation phrase. This is because he has a breathing problem which urges him to pause between segments. Let us consider the intonation of another question sentence which strengthens the above assertions.
The two productions are almost completely different in that (a) sounds purely a statement; whereas (b) sounds a genuine question. Among other things, the main reason for this difference is the final pitch patterns. One is falling and the other is rising. It is very natural for a question sentence like this one to end with a final rising intonation. The patient, however, sets the intonation the other way round.

4.4. Duration

The term duration is often associated with the time a segment takes to be produced. On the basis of this assumption, we have terms like gemmination and lengthening, to refer to consonant length and vowel length respectively. So, longer segments differ from their shorter counterparts with respect to a number of features. Among others, geminated segments are more audible and need more muscular actions than ungemminated ones. Segment duration is predictably linked to speaking rate as other acoustic qualities (e.g. formant frequencies, rate of change in formant frequencies) change when speaking rate changes.

We, in normal fluent speech, alter our rate of speech dynamically and different people speak at different intrinsic rates; as a result, the physical duration of different parts of the speech signals including the durations of segments and syllables change with a change in speaking rate. In this section, an account of segment duration in the patient’s speech will be given. Moreover, the time

---

5 There are some differences between the patient’s pronunciations of some of the words. He said /ʔihe/ to mean ‘this’, which is actually uttered as /jihe/, /ʔizi/ for /ʔizih/ glossed as ‘here’ and /misärw/ for /ʔi(jä)misärw/ which literally means ‘who works (Masculine)’. Note that these are not errors made due to his physiological problem rather they are just facts found in any person’s speech (usually in fast speech). They could also be considered as dialectal variations.
he takes to produce words, phrases and sentences in contrast with the ‘normal’ will be examined.

The patient’s speech is remarkably slow. This can be simply proved by looking at the times indicated on the spectrograms and other displays of his speech as compared to the ‘normal’. For example, the patient took 1.27063 seconds to produce the simple sentence /tát’ärät’tärku/ ‘I doubted’ while the normal person took 0.654188 seconds for the same utterance. Similarly, the patient produced another simple statement /jìmäsłáŋmal/ glossed as 'I think so' in 1.60925 seconds while the other person did it in 0.717483 seconds.

4.5. The Patient’s Duration vs. the 'Normal'

Since a segment (phoneme) is the smallest distinctive element that is capable of conveying a difference in meaning, the discussion on speech duration (speaking rate) needs to begin with segment duration. This should be carried out by looking at the time a single segment takes in different words at different occurrences (i.e., initially, medially and finally) and in longer expressions and contrasting the patient's times with what we have been referring to as 'normal'. Yet, since, in Amharic, utterances which pretend to start with vowels actually begin with the glottal stop /ʔ/, which, in other words, means that we will not examine vowel durations at the initial position—only medially and finally. In addition, there is no word that ends with /ɨ/, so we will not have value for this vowel at final position.

Disregarding individuals' variations in speech, all of the vowels of Amharic are 'short'. In other words, there is no word which is contrasted by vowel lengthening. Let us consider time differences in producing the seven Amharic vowels in the medial and final positions to see the scope of the abnormality of the patient’s vowels in this respect. In the following table “M” stands for 'Medially' and “F” for 'Finally'.

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Patient’s production</th>
<th>‘Normal’</th>
<th>Patient- ‘Normal’</th>
</tr>
</thead>
<tbody>
<tr>
<td>i</td>
<td>M 0.118197</td>
<td>M 0.099751</td>
<td>M 0.018446</td>
</tr>
<tr>
<td></td>
<td>F 0.110045</td>
<td>F 0.147210</td>
<td>F -0.037165</td>
</tr>
<tr>
<td>e</td>
<td>M 0.238560</td>
<td>M 0.117188</td>
<td>M 0.121372</td>
</tr>
<tr>
<td></td>
<td>F 0.100340</td>
<td>F 0.135375</td>
<td>F -0.035035</td>
</tr>
<tr>
<td>ɨ</td>
<td>M 0.257198</td>
<td>M 0.097527</td>
<td>M 0.159671</td>
</tr>
<tr>
<td>ä</td>
<td>M 0.120862</td>
<td>M 0.116667</td>
<td>M 0.004195</td>
</tr>
<tr>
<td></td>
<td>F 0.097582</td>
<td>F 0.139501</td>
<td>F -0.041919</td>
</tr>
<tr>
<td>ɑ</td>
<td>M 0.433873</td>
<td>M 0.157812</td>
<td>M 0.276061</td>
</tr>
</tbody>
</table>
Table 1 | Time differences (in seconds) between the patient's productions of the vowels and the 'normal' production.

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>0.014062</th>
<th>F</th>
<th>0.118118</th>
<th>F</th>
<th>-0.104056</th>
</tr>
</thead>
<tbody>
<tr>
<td>o</td>
<td>M</td>
<td>0.192772</td>
<td>M</td>
<td>0.123605</td>
<td>M</td>
<td>0.069167</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.013331</td>
<td>F</td>
<td>0.103129</td>
<td>F</td>
<td>-0.089798</td>
</tr>
<tr>
<td>u</td>
<td>M</td>
<td>0.131216</td>
<td>M</td>
<td>0.103129</td>
<td>M</td>
<td>0.028087</td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>0.014250</td>
<td>F</td>
<td>0.086644</td>
<td>F</td>
<td>-0.072394</td>
</tr>
</tbody>
</table>

Below are histograms that summarize the time differences stated in the above table.

Graph 1 | Graphic representation of the time differences between vowel productions at word medial position by the patient and the ‘normal’ person.

Graph 2 | Graphic representation of the time differences between vowel productions at word final position by the patient and the ‘normal’ person.

From Table 1 and the corresponding histograms that follow it, we can tell that the patient's speech is remarkably slow. In general, all of the patient's vowels are found to be longer than the 'normal' when they are pronounced medially. And all of them are shorter than the normal when they are pronounced finally. On average, the patient's vowels are abnormally longer by 0.096714 seconds when they occur medially; similarly they are abnormally shorter by 0.063395 seconds when they occur finally.

---

This figure does not account for /ɨ/ because this vowel does not have a value word finally.
It is true that the time that a single vowel takes might vary depending mainly on the type of sound that precedes or follows it. Considering the influence that this factor would put on the reliability of the comparison, the figures in the above table are averages of the times that every vowel (in both productions) takes when produced preceding and following different sounds. Therefore, the average times, by which the patient's vowels are abnormally longer and shorter than the 'normal' prove the claim we have been making based on mere auditory observations. One important point that should be stated here is that his vowels are even shorter and inaudible when they occur finally in words that appear at the end of a longer sentence. This may be because of the inactive muscular coordination that result from his sickness; and may also be because the energy that comes out of his lungs is minimal and thus does not last long to produce longer utterances. For this reason, he takes brief pauses between sentences and some times within a sentence to reset his energy.

As to consonant durations, we can simply see the time differences between the two productions of a word. Just to cite one instance from the data, consider the following waveforms where the word /bɨloɲ/ which would mean 'after he/it said to me' as produced by the patient (a) and by the 'normal' person (b).

![Figure 5 Waveforms of a phrase /bɨloɲ/ glossed as 'after he/it said to me' as produced by the patient (a) and the other person (b).](image)

As can be vividly seen from the waveforms, the patient took 0.678312 seconds whereas the normal person took only 0.324514 seconds to produce the same utterance. This means that the patient used, for this particular utterance, almost twice a time used by the other person. Apart from the time variation, we can also observe the waves are very much squeezed together in the patient's production but they are relatively spread out in the other person's
production. This is clearly because the same window length is used for both productions that have different durations.

Now let us look at the overall situations of the patient's speaking rate as opposed to the 'normal'. It is not surprising for different individuals to speak in various speaking rates. In spontaneous speech, some individuals speak in a fast way while others moderately and still some other people talk slowly. Although it would be difficult to label individuals' speech as 'fast', 'moderate' or 'slow' (because there are a range of speaking rates that fall in a continuum), there exists an average rate which integrates different rates of adult males' speeches.

Numerous literature provide different figures which correspond to the average speaking rate for adult male speech. However, most of the suggestions fall between 125-150 words per minute. To make use of this range as a reference scale we should be clear with what a 'word' is really supposed to mean. In languages like English where we have very simplified morphology, counting the number of words in a sentence is rather a simple task. This would not, however, be valid in languages which have complicated morphology. Amharic is one of the languages with complicated morphosyntax where a single word-like utterance can be a sentence. Therefore, trying to count the number of words uttered per minute would be difficult.

The best approach to know the problem of the patient’s speaking-rate is to look into some of his sentences and the normal person’s sentences which are of similar structures to the patient's ones and examine them from the perspective of time. By so doing, we can have a better picture of the problem of the patient's speaking-rate. In view of that, it appears that his speaking rate and fluency get slower and slower as he speaks much for an already stated reason.

4. Conclusion
In conclusion, from the discussions above, it appears that apart from the inconsistent pitch pattern that happen at the middle of the utterances, the final pitch contours that usually determine the types of sentences are not normal. As a result, it is very difficult, from the patient’s intonation patterns, to get what the patient exactly intends to mean. But it is relatively more effortless to grasp what he means from the context in which he utters sentences than in fragmented utterances like the ones just dealt with; moreover, as stated above, his speaking rate is abnormally slow.

References