DISTRIBUTION OF PITCH ACCENTS IN INTONATION GROUPS ACROSS INTONATIONAL STYLES

Dmitry Sityaev
Phonetics Laboratory, University of Oxford, Oxford, UK

ABSTRACT

The current study investigates the relation between prosodic structure and paralinguistic factors. Specifically, the number of pitch accents found in an intonation group is claimed to vary depending on an intonational style in which the speaker chooses to convey a message. It is shown that in formal speech, speakers tend to place three and four accents per intonational group more often than they do in informal speech. Also, the difference in prosodic organisation was noted between read speech and spontaneous speech. In general, it was found that in spontaneous speech, in about 50-60% of all the intonation groups speakers place only one accent per intonation group.

1. INTRODUCTION

The prosodic system of the English language has been the centre of attention for a long time now. A great deal of contribution has been made, especially in the area of intonation. Yet, descriptions of English intonation which were developed specifically for the purposes of teaching English have mainly concentrated on sentences taken in citation form [3, 4]. This seems to present no problem at the early stages of learning a language. However, problems start to arise when students attempt to produce utterances longer than just one or two sentences. It is very often the case that they still sound as if they were producing sentences in isolation. The main reason for this seems to be the fact that learners carry patterns of accent assignment which they have learnt for sentences in citation into connected speech, i.e. they produce more accents in an intonation group than native speakers normally do.

The aim of this study is to look at the level above that one of a sentence - text. Although essential rules of sentence intonology persist into fluent speech (for instance, deaccent functional words, etc.), connected speech is a different dimension and is worth while investigating independently. In this paper we aim to identify the relation between the number of pitch accents found in an intonation group and various paralinguistic factors, i.e. we examine various intonational styles as well spontaneous speech vs. read speech.

One of our goals is to demonstrate prosodic variation in different speaking styles. It has been shown that stylistic variation is responsible for intonational variation [2]. Various paralinguistic factors have been noted to influence the prosodic organisation of speech such as the subject matter (content of the text), discourse mode (spoken/written language), activity type (telling a story, giving a lecture, etc.) and the relation between participants [7]. Each of these factors seems to influence the sound pattern of speech.

Three intonational styles were chosen for the present study: academic, informational and conversational. Among other things discussed below, these three styles primarily vary in the degree of formality between the speaker and the hearer(s). Academic style is a prototype of a formal style and conversational style is a typical example of an informal style. Our claim is that the number of pitch accents per intonation group is higher in formal styles than in informal styles.

Academic style is the one which is most likely to be characterised as formal. It is frequently manifested in academic and educational lectures, scientific discussions, at conferences, seminars and in classes. The speaker usually wants to attract the listeners' attention, establish contact with the audience and direct the public attention to the message carried by the content of the speech. In most cases, public lectures are prepared in advanced which means that academic style cannot be characterised as representing spontaneous speech.

Informational style is often described as "neutral" and it is least of all affected by various extralinguistic factors. An example of the manifestation of informational style could be educational descriptive narrative, when, say, the speaker tries to tell us about some subject (be it a source-filter theory, or a story about the origin of mammoths) but does not want to sound like a lecturer. It would be wrong to call this style formal since the degree of formality between the speaker and the listener here is not the same as in the case of academic style. Yet, this style cannot be called conversational since the goal is to inform listeners of something entirely new. Thus, informational style seems to fall between academic style and conversational style.

Conversational style is the most commonly used intonation style. It is the style which is manifested in natural, spontaneous, everyday speech. Conversational style is very often referred to as informal since speakers mainly resort to it in informal settings, i.e. when they speak to relatives, friends, people whom they know quite well, etc.

We expect to find more intonation groups containing 3-4 pitch accents in academic style and more intonation groups containing 1-2 pitch accents in conversational style. This difference is expected due to several factors: mainly activity type which has an impact on the amount of information communicated and relation between the participants. Likewise, we would expect to find differences in prosodic organisation between read speech and spontaneous speech: read speech is likely to have more intonation groups with 2-3 accents in a group whereas spontaneous speech is expected to abound in intonation groups containing just one pitch accent.

2. EXPERIMENT 1

2.1. Method

2.1.1. Material and subjects. Material used for the experiment consists of recordings of spontaneous (or quasi-spontaneous in the case of academic style) English speech from three speakers.
The subjects used for this experiment were all research assistants at Oxford University aged between 24 and 28. There were 1 female speaker and 2 male speakers. All subjects were native speakers of Southern British English.

Recording was done in two stages. The first was in a classroom setting when the subject was delivering a lecture to a group of students. At the second stage, the subjects were invited to the laboratory where recordings were made for informational and conversational styles. The subjects were instructed to talk for about ten minutes firstly about the subject of their research (bearing in mind that they were talking to a non-specialist) and secondly about their previous weekend.

2.1.2. Equipment and procedure. The recordings were made in a sound treated booth (except for when subjects were giving a lecture) on DAT-tape. Speech samples were digitised at a sampling rate of 8 kHz/s on a Sun Sparcstation LX. Speech files were further analysed using ESPS/Waves+ speech analysis software.

For each file, F0 was extracted. Based on the acoustic analysis and impressionistic listening, each sample was labelled by hand for presence or absence of pitch accent using Pierrehumbert's method of description for English intonation [5]. Prosodic boundaries were marked by a vertical line1. A sample of prosodic labelling appears below in Figure 1.

If you WANT to know about MINK | MINK is a | H* !H* L-L% H* H-L%
SMALL ANIMAL | that lives by the RIVER | H* !H* L-L% L+H* L-L%

Figure 1. Prosodic labelling of a fragment of a speech sample.

2.2. Results. The results obtained for each speaker are presented in Table 1. On the basis of the data, it is possible to compare the prosodic organisation of speech belonging to academic, informational and conversational styles. All three speakers showed a similar pattern of accent distribution. Namely, intonation groups with one accent tend to be found more often in conversational style (59%-63%) than in academic style (30%-44%). Moreover, there were no four pitch accents per intonation group found for Speaker MS in informational or conversational styles, whereas for all three speakers the presence of intonation groups with three and four pitch accents in academic style is quite striking as opposed to their distribution for informational and conversational styles.

Thus, when moving from conversational style to academic style, the general tendency for all three speakers is to decrease the number of intonation groups with one accent and to increase the presence of intonation groups with three and four pitch accents.

2.3. Discussion. Of the three intonational styles analysed here, academic style appears to be somewhat different from informational and conversational styles as far as the prosodic organisation is concerned. More specifically, there is a clear tendency among the three speakers to place three and four pitch accents per intonation group more often when delivering a lecture. This can be best of all explained in the following way.

When delivering a message in academic style, the speaker desires to relate a lot of information and he achieves that by referring mainly to the world outside him. We may call this type of information “external information”, or to use a term borrowed from [1] the discourse is context-dependent. In speech it is realised in the form of information abundance which is prosodically manifested by intonation groups containing three and more words bearing accents. Contrary to that, conversational style is very likely to carry less information which is entirely new both to the speaker and the hearer; the participants structure their conversation around topics, objects, things and so on which are already familiar to them. This type of information may be referred to as “internal information”; the discourse is situation-dependent. Less new information will consequently mean fewer words which is realised in a fewer number of accented words within intonation groups (usually one or two, and very rarely three).

<table>
<thead>
<tr>
<th>Style</th>
<th>Speaker</th>
<th>One accent</th>
<th>Two accents</th>
<th>Three accents</th>
<th>Four accents</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>mean</td>
<td>S.D.</td>
<td>%</td>
</tr>
<tr>
<td>Academic</td>
<td>OM</td>
<td>43.6</td>
<td>35.3</td>
<td>6.0</td>
<td>35.2</td>
</tr>
<tr>
<td></td>
<td>MS</td>
<td>29.5</td>
<td>40.6</td>
<td>37.8</td>
<td>37.9</td>
</tr>
<tr>
<td></td>
<td>NP</td>
<td>32.8</td>
<td>40.6</td>
<td>37.8</td>
<td>37.9</td>
</tr>
<tr>
<td>Informational</td>
<td>OM</td>
<td>49.4</td>
<td>54.8</td>
<td>4.8</td>
<td>47.6</td>
</tr>
<tr>
<td></td>
<td>MS</td>
<td>53.9</td>
<td>47.6</td>
<td>39.7</td>
<td>40.0</td>
</tr>
<tr>
<td></td>
<td>NP</td>
<td>61.1</td>
<td>53.9</td>
<td>37.6</td>
<td>40.0</td>
</tr>
<tr>
<td>Conversational</td>
<td>OM</td>
<td>58.7</td>
<td>60.1</td>
<td>2.2</td>
<td>38.4</td>
</tr>
<tr>
<td></td>
<td>MS</td>
<td>63.2</td>
<td>38.4</td>
<td>34.0</td>
<td>36.3</td>
</tr>
<tr>
<td></td>
<td>NP</td>
<td>58.5</td>
<td>34.0</td>
<td>36.4</td>
<td>36.3</td>
</tr>
</tbody>
</table>

Table 1. Distribution of intonation groups by style. Number of pitch accents found in an intonation group expressed as % of all intonation for each speaker.
Another explanation of the results may follow from the theory which purports to account for the meaning that pitch accents have in the interpretation of discourse. It has been long established that accents are used to mark salience in the discourse. Particularly, pitch accents are very often associated with what can be generally termed as "new information". At present, there is very little work done on the interpretation of intonational meaning. The only serious attempt to address the issue is [6]. According to [6], pitch accents are said to concern the relationship between the propositional content of the utterances and the mutual belief between the speaker and the hearer(s) involved in the conversation. For example, the H* accent is believed to signal that the item with which this accent is associated is to be added to the hearer's beliefs. In other words, the H* accent generally conveys that the items made salient by the H* are to be treated as "new" in the discourse.

Thus, there seems to be a correlation between the intonational style chosen by the speaker to convey a message and the number of words the speaker places accent onto in his speech. If we take up the view that pitch accents are used to signal that the word is to be added to the hearer's belief space, then the results can be interpreted straightforwardly: informational abundance of a lecture will cause an increase in a number of pitch accents. And vice versa, the more conversational the style is, i.e. the more familiar the speaker and the hearer are with the objects used in the discourse, the fewer pitch accent will be found in an intonation group.

We may also conjecture that since the aim of public speaking is to attract the audience's attention, subjects may choose to accent more words to sound "more prominent". However, this factor seems to be secondary.

3. EXPERIMENT 2

3.1. Method

3.1.1. Material and subjects. The subjects chosen for the second experiment were all undergraduate students at Oxford University. There were three male speakers and three female speakers. All were native speakers of Southern British English.

3.1.2. Equipment and procedure. Subjects were given an anecdotal story which they were first asked to read to themselves. It was ensured that the subject understood the joke. Then the subject was instructed to read the story in a normal tempo. After the subject finished reading the story, the sheet would be removed from the subject and he or she would be asked to retell the story as if the subject was telling the joke to a friend.

The samples obtained were transferred from DAT-tape to Sun Sparcstation in the similar fashion as described in Experiment 1. \(F_0\) was extracted in each file and used as a guideline in prosodic labelling of the recorded material.

3.2. Results. The data obtained show the difference in the distribution of pitch accents between read speech and spontaneous speech. In spontaneous speech, speakers tend to place fewer accents per intonation group than they do when reading a passage.

The difference in distribution of pitch accents as to the number of accents found an intonation group between read speech and spontaneous speech is exemplified in Figure 2 and Figure 3. For all six subjects, there is a strong tendency to place one accent (mean = 54.6%, s.d. = 4.6) or two accents (mean = 39.1%, s.d. = 3.8) per intonation group in spontaneous speech. Read speech exhibits variety as to the distribution of pitch accents for different speakers, but by far groups with two pitch accents dominate (except for Speaker 1), especially in the cases of Speaker 3 and Speaker 6 (mean for all speakers = 52.8, s.d. = 11.8).

3.3. Discussion. Experiment 2 supports the hypothesis that the distribution of pitch accent in connected speech varies from the same sentences in citation form. Namely, in about 50-60% of all intonation groups, there is only one pitch accent in an intonation group. Intonation groups containing two pitch accents comprise around 40% of the text. Finally, three-accent intonation groups occur in less than 10% of all the intonation groups.

The tendency of having one or two pitch accents in a group in spontaneous speech can be explained by the degree of spontaneity involved. In reading the speaker has got a structured message which allows him to plan intonational phrasing, make accentual decisions and so on in advance. In spontaneous speech, the speaker is forced to think on the spot. Although the same amount of information seems to be communicated when reading or retelling a joke, prosodic organisation of read speech and spoken speech is nevertheless different.

To draw a parallel between Experiment 1 and Experiment 2, spontaneous speech can be thought of as the archetypal informal style, whereas read speech exhibits a great many features of a formal style.
4. GENERAL DISCUSSION AND CONCLUSION

The findings reveal that there exists a relation between intonational style and the number of pitch accents found per intonation group. In other words, depending on various paralinguistic factors, we may expect variation in prosodic organisation of speech (in particular the number of pitch accents found in an intonation group as shown in this study).

Since in a lecture the speaker appeals to a lot of notions which may be new for the hearer (and indeed, it is the primary purpose of the lecture - to introduce new material, to make the audience familiar with the issues, etc.), we expect to find a greater number of pitch accents found per intonation group. And vice versa, since in conversations the speaker and hearer share a lot of information, there is no need to make every object salient. Informational style may be still characterised as abundant in information; yet, its prosodic organisation puts it next to conversation style. This may be largely for the reason that the speaker tries to sound informal and explain things in a way that would be easy for the hearer to understand, i.e. resorting to things which are already part of our knowledge.

The results of the present study also suggest that in spontaneous speech, there is a strong tendency to find only one pitch accent in an intonation phrase. This observation may have the following implications.

Firstly, the framework of intonation treatment proposed by Pierrehumbert [5] is adequate in the sense that it does not differentiate between the traditional notions of 'pre-nuclear accent' and 'nucleus'. As it has been shown above, in about 50-60% of all intonation groups there is only one pitch accent which raises the question whether we should still recognise intonational heads proposed in [4].

Secondly, the percentage of intonation groups with one, two and three pitch accents found in fluent speech may be taken into account by those working in the area of speech synthesis. The current study has shown that generally it is one or two accents that are found per intonation group. The rest material in a group appears to be deaccented. The abundance of pitch accents will in no way increase the naturalness of synthesised speech. Or so it seems at least given the premise that the idea behind speech modelling lies in maximally approximating the mechanisms employed by machines to those ones used by humans.

And lastly, those who are engaged in language teaching may take into account the fact that the number of pitch accents in one or two. This is especially important to point out to students who aim at achieving fluency in English. In spontaneous speech, many words surface as de-accented. The focus seems to be on working out where sentence accent is more likely to be placed. It is believed that accentual decisions are made with respect to the information structure and argument structure of a sentence, but more research is needed in this area.

ACKNOWLEDGEMENTS

I would like to thank Andrew Slater for his help with recording and analysing speech samples.

NOTES

1. For the purposes of the present study, it was assumed that there is no level of intermediate phrase, i.e. speech contains only full intonation phrases. Consequently, the phrasal tone L- or H- is used only to mark the changes in F0 contour and does not mark any boundaries.