Casual speech tends to exhibit a larger proportion of reduced sound segments than careful speech; e.g. she packs shorts, will be /ʃi:pks/ rather than /ʃi: pa:kz/. In the present paper, we ask whether we can find comparable effects in English intonation.

Pitch accent distribution in three speaking styles was examined: read, semi-spontaneous and spontaneous speech. The data were taken from the IViE corpus [4], and were transcribed prosodically in the IViE labelling system, an autosegmental-metrical transcription system for comparative intonation analysis. After labelling, pitch accent types were counted and compared. The results showed that in casual speech, speakers used fewer complex pitch accents than in careful speech. Our findings offer support for autosegmental approaches to intonation analysis such as the ones proposed by [5] and [2] which allow for a number of phonological processes in intonation, including continuous speech processes.

1. INTRODUCTION

In segmental phonetics, differences between careful and casual speech have been modelled by continuous speech processes such as assimilation or elision [6], and it is generally assumed that casual speech exhibits more instances of such processes than careful speech. In the present paper, we ask whether we can find comparable evidence for continuous speech processes in English intonation.

Direct analogies between segmental and prosodic levels of linguistic structure can be derived from the autosegmental-metrical models of intonation proposed by Gussenhoven [5], and Grabe [2]. These models postulate a limited set of tonal alteration rules which modify the structure of pitch accents in certain contexts, and a subset of these tonal alteration rules can be compared to continuous speech processes. For instance, in prenuclear position, a falling accent H*_+L, may undergo DISPLACEMENT [2], [5]’s ‘partial linking’), a process which is similar to regressive assimilation. In regressive assimilation, a postulated segment such as the /s/ in /ʃi: pa:kz/ is modelled as having adopted some or all of the characteristics of the following segment /ʃi:/, and one may say that as a result, the assimilator and the assimilee have become more similar, or more closely related structurally. When DISPLACEMENT has applied, a comparable process has taken place; in certain contexts, the second tone of a pitch accent is shifted away from the first to the right, approaching more closely the following pitch accent. In Grabe’s account, H*_+L with DISPLACEMENT is transcribed as H*_+L, and this is the notation we will use in what follows. In H*_+L, the fall in pitch is more gradual than in H*_+L, and spans several syllables. Figure 1 below illustrates H*_+L without DISPLACEMENT on the left, and with DISPLACEMENT on the right.

In the present study, we hypothesised that DISPLACEMENT and DELETION can be compared to continuous speech processes such as assimilation and elision. Our hypothesis predicts that we should find a higher proportion of pitch accents which have undergone DISPLACEMENT and DELETION in faster, more casual speech, than in careful speech. Our experimental methodology is described in the following section.

2. THE EXPERIMENT

2.1. Method

2.1.1. Data
We analysed a subset of data from the ‘Intonation in the British Isles’ corpus [3, 4]. The data were produced by two speakers of General Southern British English (GSB) in three speaking styles: read speech, semi-spontaneous speech, and free conversation. We assumed that the degree of formality would decrease from one speaking style to another; the read speech would be the most formal, the semi-spontaneous speech would be intermediate, and the spontaneous speech would be the least formal. From each speaking style, we analysed about 10 minutes of speech. Data from GSB, rather than data from another variety of English was chosen for the purposes of the present study because the Gussenhoven system, the first
derivational account of English intonation, was designed for Southern Standard British English.

2.1.2. Subjects
Data from one male and one female subject were analysed. The speakers were pupils at Hills Road Sixth Form College in Cambridge and were 17 years of age. They had always lived in the Cambridge area, and their parents came from the South of England. The spontaneous speech data were produced in conversation with a further two speakers, one male and one female (single-sex pairs). The data produced by the additional speakers in the free conversation was included in the comparisons discussed below.

2.1.3. Tasks
The data were elicited as follows.

(a) Read speech
Speakers were asked to read aloud a modified version of the fairy tale ‘Cinderella’. Modifications involved, for instance, the replacement of words with little voiced segmental materials with others which contained more voicing, the addition of, and a range of tags. ‘Cinderella’ was chosen, because this particular fairy tale is widely known, and as a result, the possible interpretations of the text are more limited than if speakers had been asked to read a text new to them. Additionally, Cinderella is commonly read to children, and the text is therefore likely to elicit animated and varied productions with large pitch excursions. Such data offer a felicitous starting point for prosodic analysis, especially when supplemented with comparable data from other speaking styles.

(b) Semi-spontaneous speech
After having read the passage of text, the speakers were asked to retell the fairy tale from a series of pictures. No time was given for preparation, the speakers were not prompted at any time, and they were not allowed to make notes before they started.

(c) Spontaneous speech
In the third task, the speakers were paired with another interviewee, and were asked to talk for about 5 minutes about the effects of tobacco advertising on their view of smoking (the topic had been in the news at the time the recordings were made).

2.2. Analysis
The data were transcribed by the first author in the IViE labelling system (/avíː/, Intonation Variation in English [3], IViE is an autosegmental-metrical transcription system, designed specifically for comparable transcriptions of data from different varieties of one language, or different speaking styles. IViE works in conjunction with Xwaves (Entropic) under UNIX. The IViE transcriber tool is modelled on the ToBI labelling tool [1], and shows a waveform and the corresponding fundamental frequency (F0) trace, as well as a labelling template (but note that IViE and ToBI labels differ, and that IViE has a larger number of tiers for transcription than ToBI). The IViE labelling template has five tiers on which the orthography, the rhythmic structure, the auditory phonetic realisation and the phonological structure of texts can be transcribed. In the present paper, our discussion is limited to the phonological tier.

After labelling, a subset of the transcriptions was discussed and verified with three other trained transcribers, and in a small number of cases the labelling was adjusted (majority decisions were taken). Then, the distribution of a subset of labels was examined in each speaking style. This subset of labels involved three types of pitch accents:

(1) Tone labels transcribing fully realised accents
H*+L high target followed by low target
L*+H low target followed by high target

(2) Tone labels involving DISPLACEMENT
H*+ L low target shifted to the right
L*+ H high target shifted to the right

(3) Tone labels involving DELETION
H* high pitch target, low target deleted
L* low target, high target deleted

The text files containing the phonological labels were transferred to a spreadsheet (Microsoft Excel), tone labels other than the ones listed under (1)-(3) above were deleted (e.g. boundary tones, or labels for rise-falls). Then the data were separated into two groups:

Group I: All pitch accents
Group I contained labels for all pitch accents listed under (1)–(3) above separately for each speaking style. The instances of each accent type were counted with the help of a script written in Perl, and sorted into the categories ‘speaking style’ (read, semi-spontaneous, spontaneous) and ‘complexity of pitch accent’ (full, displaced, deleted) in Excel.

Group II: Prenuclear pitch accents
Group II contained labels for prenuclear accents, again separately for each speaking style. The data were counted and sorted in the same way as the data in Group I above.

Prenuclear accents were examined separately because the models proposed by [2] and [5] predict that in English, tonal alteration rules such as DISPLACEMENT and DELETION apply to prenuclear, but not nuclear accents ([2] suggests that the application and function of tonal alterations is language-specific, and shows that in German, tone DELETION applies to prenuclear and to nuclear accents).

2.1.1. Predictions
We predicted that instances of tone DISPLACEMENT and tone DELETION should increase from read speech to free conversation.

2.3. Results
2.3.1. Group I: All accents
Table 1 below shows the raw results for all accents and Table 2 gives the percentage occurrence of a particular accent type in each speaking style (in Table 2, the findings which support our predictions are given in bold). Table 1 shows that 1275 pitch accents were labelled in all; 627 in the read speech data, 399 in the semi-spontaneous speech, and 249 in the free conversation.
Table 1. All pitch accents: raw data.

<table>
<thead>
<tr>
<th></th>
<th>Full</th>
<th>DISPLACE-MENT</th>
<th>DELETION</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read speech</td>
<td>453</td>
<td>35</td>
<td>139</td>
<td>627</td>
</tr>
<tr>
<td>Semi-spont.</td>
<td>232</td>
<td>80</td>
<td>87</td>
<td>399</td>
</tr>
<tr>
<td>Spont. speech</td>
<td>132</td>
<td>19</td>
<td>98</td>
<td>249</td>
</tr>
</tbody>
</table>

Table 2 shows that the percentage of fully realised accents decreased from read speech to spontaneous speech, even when we combine nuclear and prenuclear accent patterns, whereas the number of accents with DELETION increased.

<table>
<thead>
<tr>
<th></th>
<th>Full</th>
<th>DISPLACE-MENT</th>
<th>DELETION</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read speech</td>
<td>72%</td>
<td>6%</td>
<td>22%</td>
<td>100%</td>
</tr>
<tr>
<td>Semi-spont.</td>
<td>58%</td>
<td>20%</td>
<td>22%</td>
<td>100%</td>
</tr>
<tr>
<td>Spont. speech</td>
<td>53%</td>
<td>8%</td>
<td>39%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 2. All pitch accents: percentages.

Table 2 also shows that the results for our intermediate category 'DISPLACEMENT', however, do not appear to conform with any pattern and do not appear to support our predictions. However, as DISPLACEMENT and DELETION were predicted to characterise accents in prenuclear position only, we need to examine the labels for prenuclear accents separately.

2.3.2. Results for Group II: prenuclear accents.

The results support our predictions; speakers produced more fully realised prenuclear accents in read speech than in semi-spontaneous speech than in the free conversation.

The results for accents with DELETION supported our predictions and are illustrated in Figure 4. Figure 4 shows that we found more instances of DELETIONS in semi-spontaneous and in spontaneous speech than in read speech. Note, however, that the results for read and semi-spontaneous speech do not seem to differ, suggesting that with respect to DELETION, read speech as a speaking style may differ crucially from spontaneously or semi-spontaneously produced speech data.

Table 3 below gives the raw data for the three pitch accents types in the three speaking styles with percentages.
reflect the application of a phonological process in intonation, second target in a bitonal pitch accent to the right) may not and auditory effect of DISPLACEMENT (i.e. the shifting of the DISPLACEMENT may require some rethinking. The acoustic DELETION supported our prediction. In read speech, we in speaking style. This findings suggests that the status of undergone DISPLACEMENTS did not correlate with changes in speaking style. Our findings for fully realised accents and accents with DELETION, The number of accents labelled as having more reduced pitch accents than read speech.

At the beginning of the present paper, we predicted that more formal speaking styles would be characterised by a larger number of complex accents that less formal speaking styles. Our findings for fully realised accents and accents with DELETION supported our prediction. In read speech, we found more complex, fully realised pitch accents than in semi-spontaneous speech and free conversation. Conversely, the semi-spontaneous speech and the free conversation exhibited more reduced pitch accents than read speech.

Our findings can be accounted for parsimoniously in 'derivational' models of English intonation such as the ones proposed by [5] and [2] which allow us to account for segmental and suprasegmental reduction effects with similar mechanisms. In [5] and [2], the assumption is that in certain contexts, pitch accents are subject to a limited set of intonological adjustments, and [2] hypothesises that one context which is likely to trigger the application of the tonal reduction processes DEPLACEMENT and DELETION is a change in speaking style. Our result for DELETION support this hypothesis; when the speaking style becomes less formal, we find more instances of tonal reduction. Non-derivational models of intonation which do not allow us to draw explicit analogies between segmental and suprasegmental levels of represenatation (e.g. the system which underlies [1]), cannot capture segmental and suprasegmental reduction effects with comparable mechanisms.

Secondly, our results suggest that the nature of DISPLACEMENT differs from that of DELETION. Unlike fully realised pitch accents and pitch accents with DELELTION, The number of accents labelled as having undergone DISPLACEMENTS did not correlate with changes in speaking style. This findings suggests that the status of DISPLACEMENT may require some rethinking. The acoustic and auditory effect of DISPLACEMENT (i.e. the shifting of the second target in a bitonal pitch accent to the right) may not reflect the application of a phonological process in intonation, but a difference in the phonetic realisation of an intonation pattern. Future research in this area could involve a perceptual experiment comparing the perceptual difference between fully realised accents and pitch accents with DISPLACEMENT and DELETION. The findings from the present paper suggest that listeners may judge that the difference between fully realised and deleted versions of H*+L or L*+H is significantly larger than the difference between fully realised accents and accents with DISPLACEMENT.

5. SUMMARY AND CONCLUSION
In the present paper, we investigate pitch accent distribution in three speaking styles: read speech, semi-spontaneous speech, and free conversation. Our results suggest a correlation between pitch accent complexity and speaking style; more formal speech is characterised by a larger number complex pitch accents (e.g. H*+L), and less formal speech is characterised by a larger number of simplex, 'reduced' pitch accents (e.g. H*). We argue that our findings illustrate an analogy between segmental and suprasegmental representations of speech, and that they lend support to derivational models of intonation such as the ones proposed by [5] and [2]. The models proposed by [5] and [2] offer mechanisms which can capture parsimoniously the analogy between between segmental and suprasegmental levels of linguistics structure which we characterise our data. In segmental phonetics, reductions of segmental structure have been associated with more casual speaking styles, and have been modelled as phonological processes such as assimilation or elision. Our findings suggests that similar effects apply in intonation. The models put forward by [2] and [5] allow us to capture segmental and suprasegmental reduction effects with comparable mechanisms at different levels of linguistic structure.

Table 3. Raw data for prenuclear accents.

<table>
<thead>
<tr>
<th></th>
<th>Full</th>
<th>DISPLACE-</th>
<th>DELETION</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read</td>
<td>199</td>
<td>29</td>
<td>138</td>
<td>366</td>
</tr>
<tr>
<td>speech</td>
<td>(54%)</td>
<td>(8%)</td>
<td>(38%)</td>
<td>(100%)</td>
</tr>
<tr>
<td>Semi-</td>
<td>56</td>
<td>11</td>
<td>89</td>
<td>156</td>
</tr>
<tr>
<td>spont.</td>
<td>(36%)</td>
<td>(7%)</td>
<td>(57%)</td>
<td>(100%)</td>
</tr>
<tr>
<td>speech</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spont.</td>
<td>50</td>
<td>16</td>
<td>86</td>
<td>152</td>
</tr>
<tr>
<td>speech</td>
<td>(31%)</td>
<td>(11%)</td>
<td>(58%)</td>
<td>(100%)</td>
</tr>
</tbody>
</table>

ACKNOWLEDGEMENTS
The work reported in this paper was supported by a grant from the Economic and Social Research Council to F. Nolan and E. Grabe, ESRC grant R000237145, 'Intonation in the British Isles'. Also, we are grateful to Eric Fixmer for help with Perl.

REFERENCES