1. Introduction
The production of English utterances by Polish students results in systematic replacement of English sounds by what is felt to be their Polish equivalents. These substitutions add up to make a strongly accented variety of English known as ‘Polglish’ (Sobkowiak 2001). A number of recent publications that come from both native English and foreign quarters determine the goals to be reached in the pronunciation of English by foreigners. Yet studies enquiring into the perceptual judgements of accented speech by native speakers of English are scarce. The present paper aims at filling this lacuna by determining a hierarchy of Polish accent features that strike the English ear. The study was based on the recordings of English texts read by 122 Polish university students.

2. Material
Two types of utterances were recorded – read passages and spontaneous conversation with the experimenter. The recordings were made at 7 institutions of higher education, both university departments of English (Lublin, Warsaw, Wrocław) and teacher training colleges (Kielce, Krosno, Rzeszów). The students came from rather extensive areas surrounding the schools, so the project covers speech representative of 1/6 of Poland’s area.
In total, the material comprised over 20 hours of recording. The recordings were made on a SONY MDLP minidisk with a SONY ECM-MS907 microphone, which produced satisfactory quality. The text used was the “Diagnostic Passage” from Celce-Murcia et al (1996: 398), quoted below:

If English is not your first language, people may have noticed that you come from another country because of your “foreign accent.” Why do people usually have an accent when they speak a second language?

Several theories address this issue. Many people believe that only young children can learn a second language without an accent, but applied linguists have reported cases of older individuals who have mastered a second language without an accent. Another common belief is that your first language influences your pronunciation in a second language.

Most native speakers of English can, for example, recognise people from France by their French accents. They may also be able to identify Spanish or Arabic speakers over the telephone, just by listening carefully to their pronunciation. Does this mean that accents can’t be changed? Not at all! But old habits won’t change without a lot of hard work, will they?

In the end, the path to learning to speak a second language without an accent appears to be a combination of hard work, a good ear, and a strong desire to sound like a native speaker. You also need accurate information about the English sound system and lots of exposure to the spoken language. Will you manage to make progress, or will you just give up? Only time will tell, I’m afraid. Good luck, and don’t forget to work hard!

The proportions of occurrence of sounds and classes of sounds: are shown in the illustration below:
3. Objective error count

The recordings were evaluated to allow the choice of material for the audio tape. The evaluating panel consisting of 12 independent judges listened to 122 recordings of Polish students. Each text was confronted with its exemplary transcription elaborated before to identify and mark every pronunciation error. The results were normalized to 100%. A two-way ANOVA determined that the scores for Error Types and Recordings differed significantly at 0.001, cf. Table 1 below:

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
<th>Test F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recordings</td>
<td>172623,3</td>
<td>121</td>
<td>1426,639</td>
<td>3.906581</td>
<td>5.64233E-36</td>
<td>1.470227</td>
</tr>
<tr>
<td>Error Types</td>
<td>346130,4</td>
<td>13</td>
<td>26625,41</td>
<td>72.90867</td>
<td>1.1508E-150</td>
<td>2.675961</td>
</tr>
<tr>
<td>Error</td>
<td>574441,7</td>
<td>1573</td>
<td>365,1886</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1093195</td>
<td>1707</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. ANOVA Results at α=0.001

Figure 2. Proportion of studied phenomena in the text used in the experiment
Errors made while reading were first evaluated by a team of Polish expert phoneticians. It was found that three Error Types were committed in 40-55% of tokens; these were: rhotacization of vowels (55% of cases), incorrect rendering of the interdental fricatives (48% of cases), and complete devoicing of the lenis obstruents (42%).

The occurrence of errors was confirmed by acoustic analysis. Let us now consider articulatory characteristics and acoustic reflections of these three errors.

3.1. Rhotacization

Rhotacized vowels are articulated with the front part of the tongue raised towards the post-alveolar region, which manifests itself in the lowering of $F_2$ and $F_3$ and a very strong weakening (sometimes disappearance) of $F_4$. The dip in the formants is of course inversely related to the degree of raising of the tongue front. In Polish-accented rhotacization of the unchecked vowels /a:/, /ɔ:/, and /ɔ:/ and the centring diphthongs /ɔə/ /ɛə/ and /uə/ it often happens that the rhotic modification is delayed until 1/3 of the vowel has been pronounced, and then it gradually increases, reaches its peak and descends, being reflected in the dip+rise shape of $F_2$ and $F_3$ bends; cf. Figure 3:

![Figure 3. Spectrogram and oscillogram of ‘bird’ pronounced as [bɔːd] and [bɔːd]](image-url)
3.2 [θ] [Ω] - Fronting

The R.P. interdental fricatives are often replaced by Polish phonemes which are considered perceptually ‘similar’: [f], [t] or [s] for the voiceless series, and [v], [d] [dz] and [z] for the voiced series. It is relatively easy to distinguish the mispronunciation of [θ] from [t] or [s], while often the distinction between it and [f] (and [δ] vs. [v]) remains quite difficult to discern. Let us now consider the spectral difference between the interdental and labio-dental fricatives:

The coronal articulation of [θ] in which the lips do not take part allows them to round more and extend the target element of the diphthong (frame in Fig. 4) than in the decoronalised version ([f] replacement) when the lips have less time to round as they anticipate the labio-dental contact. This is clearly seen in an F2 rise in the [au]-[θ] transition, as opposed to the lack of this F2 rise in the non-coronal [f].

Secondly, the non-strident [θ] has a relatively smooth spectral structure with high energy at low and high frequencies (40 and 30 dB, respectively, with a trough reaching as low as 10 dB), while the strident [f] has conspicuous concentrations of energy between 1 and 3, and at 3,6 and 4,5 kHz; cf. the spectra in Fig. 5.
3.3 Voicing in word final lenis plosives

Voicing in stop clusters is relatively easy to observe both on the frequency and time axes. On the former, it is manifested by the occurrence of the ‘voice bar’ at the bottom of the spectrogram that runs for 120 out of 150 ms. of the closure, while on the latter it is to be seen in the gradually decreasing envelope of the phonation record; cf. Figure 6.

4. Correlations between variables

In the next stage of the analysis, correlations between all possible pairs of variables were studied. The results confirmed the validity of conclusions concerning the relative frequency of the data since the presence of vowel rhotacization and incorrect rendering of interdentals exhibited a significant (99.9%) correlation with at least two other factors:
The concept of correlation as defined mathematically refers to co-occurrence rather than to a cause-and-effect relationship. That is, we do not imply that a rhotacization error is the cause of mispronunciation of mid central vowels or wrong rendering of palato-alveolars, or that it causes wrong stress patterns. Similarly, it is not claimed that the mispronunciation of interdentals causes errors in the pronunciation of mid central vowels or palato-alveolars.

There is, however, an intricate relationship between some of the variables so that the frequent co-occurrence of some pairs can be accounted for on strictly phonetic grounds, having a cause in “the tongue” (articulation) or in “the ear” (perception). We shall discuss below those instances of the co-occurrence of variables that we argue are substantially related. There are three such explicable correlations that will be taken one by one below, while the others are instances of co-occurrence that cannot be explained on substantial grounds.

4.1. Correlation 1: Rhotacization and mispronunciation of the mid-central vowel

The co-occurrence of erroneous rhotacization of vowels is schematically represented in Fig. 8, where the class of frequently rhotacised monophthongs (horizontal filling) intersects with the class of mid-central vowels (vertical filling). Thus the correlation is based on an intersection of two variables, and refers to its overlapping part.
4.2. Correlation 2: Rhotacization and word stress

The correlation of rhotacization with word stress can be explained by linking it to the correlation described previously, i.e. if the agentive ending –er in ‘teacher’ is erroneously for BE rendered as [ər], it attracts more word stress than in the correct pronunciation in which the word final [ə] is extremely weak, contributing to the overall impression of overstressed pronunciation.

4.3. Correlation 3: Interdentals and palato-alveolars

This correlation can be explained straightforwardly, if these two factors are considered jointly under the caption ‘place of articulation’. The existence of the correlation implies that the speakers who have problems with the articulation of interdentals are also unable to pronounce palato-alveolars.

5. Error evaluation by Englishmen and Polish teachers of English

An arithmetical procedure described elsewhere (Gonet and Pietroń, in print) allowed us to find passages representative of each given Error Type that contained at least 50% of potential errors and relatively few representatives of other Error Types. Each of 10 Error Types was represented by 3 sentences.

At this point the experimental recording material was produced. Three sentences for each Error Type were provided, so that each Error Type would be represented in equal proportions. A questionnaire for native speakers was worked out, on which the 33 judges at University College London were to evaluate each of the 30 items on a 5-point scale, where a low mark meant a departure from the RP pronunciation norm, and a high score meant close resemblance to, or identity with,
RP norms. The judges were mostly students and staff at the Department of Phonetics and Linguistics and Department of Psychology, although persons not associated with the UCL were also asked to participate. They were asked to evaluate the sentences based on the features that struck their ears, and no distinction between foreign accent and comprehension was made. We have relied on the assumption that there is a straightforward one-way relationship between intelligibility and foreign accent in that what is unintelligible certainly contributes to the overall impression of the foreign accent. The reverse is not necessarily true, because there are features responsible for foreign accent that do not impede intelligibility. Although we did not deal with intelligibility *per se*, the lack of intelligibility was naturally included as adding to a foreign sound of English.

During the presentation session, the sentences were replayed from a SONY MDLP minidisk over a set of stereo speakers in a normal acoustics room to groups of 3-5 listeners. The recording consisted of three sets of ten sentences spaced with 10-sec. intervals, with the three sets being followed each by a one-minute break. Below a replica of the questionnaire is shown:
Figure 9. A replica of the Questionnaire used in the present experiment – the respondent, S. P., is a Speech Sciences student.

The same perception test was conducted for a group of Polish teachers of English. The table below shows a comparison of the results of judgements by these two groups of evaluators; the scores are standardized, with the maximum
being 100; the higher the score, the stronger negative effect is exerted by a given factor.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Evaluation by Poles</th>
<th>Score</th>
<th>Rank</th>
<th>Evaluation by Englishmen</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Interdental s</td>
<td>74</td>
<td>1</td>
<td>Interdental s</td>
<td>64</td>
</tr>
<tr>
<td>2</td>
<td>Vowel quality</td>
<td>60</td>
<td>2</td>
<td>Velar nasal</td>
<td>62</td>
</tr>
<tr>
<td>3</td>
<td>Velar nasal</td>
<td>59</td>
<td>3</td>
<td>Obstruent devoicing</td>
<td>54</td>
</tr>
<tr>
<td>4</td>
<td>Aspiration</td>
<td>57</td>
<td>4</td>
<td>Place of articulation</td>
<td>54</td>
</tr>
<tr>
<td>5</td>
<td>Place of articulation</td>
<td>57</td>
<td>5</td>
<td>Aspiration</td>
<td>52</td>
</tr>
<tr>
<td>6</td>
<td>Obstruent devoicing</td>
<td>50</td>
<td>6</td>
<td>Vowel quality</td>
<td>46</td>
</tr>
<tr>
<td>7</td>
<td>Rhotacisation</td>
<td>49</td>
<td>7</td>
<td>Obstruent voicing</td>
<td>42</td>
</tr>
<tr>
<td>8</td>
<td>Prosody</td>
<td>42</td>
<td>8</td>
<td>Prosody</td>
<td>42</td>
</tr>
<tr>
<td>9</td>
<td>Obstruent voicing</td>
<td>41</td>
<td>9</td>
<td>Rhotacisation</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>Linking r</td>
<td>12</td>
<td>10</td>
<td>Linking r</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 2. Evaluation of texts by Poles and Englishmen: results

The results obtained exhibit a very high Pearson correlation coefficient obtaining between Polish and English judgements, viz. r=0,899 which is significant at 99,9%. Both groups of evaluators agree that the mispronunciation of the interdental fricatives and the velar nasal is very annoying. An incorrect rendering of vowels adds more to the foreign accent in the opinion of Polish teachers of English, while obstruent devoicing, with associated vowel clipping, fares high according to native English evaluators. The lack of aspiration and an incorrect rendering of place of articulation win a medium rank. Other indicators of Polish accent include faulty prosody and missing linking r’s. The results in Table 2 can be now collapsed to show aggregated results in the form of a hierarchy of importance; this is done in Fig. 10.

Figure 10. Hierarchy of foreign accent errors determined on the basis of results aggregated for English and Polish judges
There is an interesting relationship between the rank placement of Error Types in the hierarchy (plot in Fig. 10), and their frequency of occurrence in the text (Fig. 2). It can be noticed that the Error Types placed at the top of the hierarchy are relatively infrequent in the text.

6. Error Types determined by native speakers’ verbal commentary

The Questionnaire filling session was followed by a discussion with the native speakers, based on the notes they had taken down while filling in the questionnaire; portions of the tape were replayed if necessary, and a leisurely atmosphere encouraged relaxed interaction. It emerged from the discussion that the disturbances in the suprasegmental structure of the utterances were very annoying.

The results gained from the questionnaires and from the discussion are highly correlated ($r=-0.7$) despite the fact that rank placements of some features vary. The Error Types that were singled out as most striking to native English ears are shown in Figure 11:

![Figure 11 Frequency of errors singled out in discussion](image)

Three Error Types fared high in both sources of information, i.e. in the questionnaires and in the verbal commentary: (1) incorrect vowel quality (which
includes rhotacization); (2) mispronunciation of the interdental fricative, and (3) incorrect devoicing of word final voiced obstruents.

The native speakers’ verbal judgements were not restricted by the implicit suggestions that every questionnaire has, and they broadened the range of Error Types with mispronunciations of the lateral – an English sound too often taken to be the same as the Polish /l/. The identification of /l/ as a possible source of foreign accent is understandable with regard to ‘dark /l/’, alien to most speakers of Polish, but quite unexpected with regard to ‘clear /l/’, usually taken to be the same in the two languages. Instances of mispronounced syllabic dark [l]’s were found to be particularly annoying due to the lack of back vowel resonance and to presence of full devoicing; cf. the spectrograms in Fig. 12.

![Figure 12. Correct (left) and incorrect (right) realization of /people/](image)

Another range of phenomena evaluated as factors important for the origination of the foreign accent comprised suprasegmental phenomena. The native users of English pointed out that, despite the relatively acceptable phonemic structure of the utterances, intelligibility was frequently hindered by wrong stressing which resulted in wrong rhythm, accompanied by inappropriate intonation.

7. Checking the substantiality of native speakers’ judgements

The choice of the sentences for the audio presentation was done according to the arithmetical determination of the occurrence of Error Types. However, there exists a danger that the English listeners gave grades based on other discriminants buried in the experimental setup and not conspicuous to the authors. To verify the
substantiality of native judgements, the native English evaluations of pronunciation were juxtaposed with the results of objective error count by Polish experts (cf. 2 above), and correlations of these two arrays were established. It was found that Pearson r reached very high values shown in Fig. 13.

![Figure 13. Pearson r values for perceptually determined Error Types confirmed by objective analysis with critical values of r significant at 80% for place of articulation, 95% for interdentals, and 99% for vowel quality and prosody.]

8. General conclusions

There are two other important studies of the perception of foreign accent in Polish learners of English; let us juxtapose their findings with ours:

<table>
<thead>
<tr>
<th>Scheuer</th>
<th>Szyra-Kozłowska</th>
<th>G&amp;P Experiment</th>
<th>G&amp;P Discussion</th>
<th>G&amp;P Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td>ET Voicing</td>
<td>Interdentals</td>
<td>Interdentals</td>
<td>Vowel Quality</td>
<td>Prosody</td>
</tr>
<tr>
<td>ET Devocing</td>
<td>Spelling pronunciation</td>
<td>Velar nasal</td>
<td>Lateral</td>
<td>Vowel quality</td>
</tr>
<tr>
<td>ET /l/</td>
<td>Devocing</td>
<td>Place of articulation</td>
<td>Devocing</td>
<td>Interdentals</td>
</tr>
<tr>
<td>NS /i/</td>
<td>word stress</td>
<td>Aspiration</td>
<td>Intonation</td>
<td>Place of articulation</td>
</tr>
<tr>
<td>P Interdentals</td>
<td>/I/</td>
<td>Vowel quality</td>
<td>Interdentals</td>
<td></td>
</tr>
<tr>
<td>P schwas</td>
<td>Trilled /t/</td>
<td>Devocing</td>
<td>Rhythm</td>
<td></td>
</tr>
<tr>
<td>ET schwas</td>
<td>-ing</td>
<td>Prosody</td>
<td>Sound duration</td>
<td></td>
</tr>
<tr>
<td>NS schwas</td>
<td>individual words</td>
<td>Obstr. Voicing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P vowels</td>
<td>Voice assimilation</td>
<td>Rhotacisation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NS vowels</td>
<td>Long and short vowels</td>
<td>Linking r</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Comparison of findings in Scheuer (2000), Szyra-Kozłowska (this volume) and this paper (G&P). Abbreviations in Scheuer are ET: ‘Native Teachers of English’, ‘NS’ – native speakers, ‘P’ – Polish teachers of English
Scheuer (2000) focuses on three major categories: (i) voicing, comprising the incorrect *devoicing of voiced obstruents* and the incorrect *voicing of voiceless obstruents*; (ii) the incorrect rendering of *interdentals*, and (iii) various aspects of mispronunciations of *vowels* (failure to pronounce the schwa vowel, the non-existence of the tense-lax contrast within the high front vowels, and mispronunciation of ‘other vowels’). It is interesting that for native teachers of English the incorrect voicing of voiceless obstruents should have fared higher than incorrect devoicing because it is a feature less frequent than devoicing. This finding has not been confirmed in any other study.

The incorrect articulation of the interdentals was found striking to the ears of Polish teachers of English, while it did not stand out for the native speakers (!), while both in Szpyra-Kozłowska’s study and ours it was rated very high.

The category that Scheuer calls ‘non-schwas’ can comprise a number of different phenomena such as, e.g. the lack of weak articulation of a vowel *[aidia], *[tiːtʃa], *[neiʃen], or incorrect rhotacization (‘r-colouring’) of certain vowels where there is the letter {r} in spelling, viz. *[tiːtʃɔr].

Besides the canonical foreign accent errors such as the mispronunciation of the *interdentals*, the lack of the *[iː] – *[t] contrast, and the *devoicing of voiced obstruents*, Szpyra-Kozłowska includes in her study the behaviour of obstruents in clusters (*voicing assimilations*), and broadens the range of foreign accent phenomena with *trilled r*, the pronunciation of the *-ing* ending and word stress. Besides these ‘atomic’ categories, Szpyra-Kozłowska also takes into account ‘spelling pronunciation’ and ‘individual words’ which can sometimes overlap with the ‘atomic’ error types.

In the present study, we have looked at the problem of foreign accent from three separate perspectives: (a) objective judgements by English and Polish judges; (b) discussion with English listeners and (c) studying the correlations of objective error counts with experimental native speaker opinions. Besides the rendering of the interdental fricatives, voicing, and vowel quality understood in a broad sense, we have taken into account aspiration and place of articulation, and a range of suprasegmental phenomena (word accent, rhythm, intonation). A number of error types involving seemingly inoffensive substitutions surfaced as important.
to the native speakers, viz. the mispronunciation of the English lateral and the failure to use the English context-conditioned sound duration variability.

It is rather difficult to compare the details of the findings in the sources mentioned above because of differences in the methods used in the study, the content of the Error Type categories, and the scales of measurement applied in the experiment. Despite this, quite a clear picture emerges with regard to error types responsible for the native speakers’ impression of ‘Polish accent’. Thus, in all the discussed studies, there are three Error Types that recur: (1) the incorrect pronunciation of the interdental fricative [θ], (2) the incorrect rendering of voicing, and (3) mispronunciations of vowel quality (most notably, of the front high vowel contrast between [i:] and [ɨ]). In our study, two other Error Types emerged as important, viz. incorrect place of articulation (palato-alveolars and alveolars) and prosody, comprising word stress, rhythm and intonation.

These findings should attract the attention of curricula-forming bodies and determine the scope of pronunciation teaching at schools and universities.

* This study has been done within a project financed by the Committee for Scientific Research in the years 2003-2006.

References


Szpyra-Kozłowska, J. (this volume). Intelligibility versus Polish accent in English.