Chapter 9

Automatically identifying characteristic features of non-native English accents

Jelke Bloem
Amsterdam Center for Language and Communication, University of Amsterdam

Martijn Wieling
Center for Language and Cognition, University of Groningen

John Nerbonne
Center for Language and Cognition, University of Groningen, Freiburg Institute for Advanced Studies, University of Freiburg

We demonstrate the application of statistical measures from dialectometry to the study of accented English speech. This new methodology enables a more quantitative approach to the study of accents. Studies on spoken dialect data have shown that a combination of representativeness (the difference between pronunciations within the language variety is small) and distinctiveness (the difference between pronunciations inside and outside the variety is large) is a good way to identify characteristic features of a language variety. We applied this method from dialectology to transcriptions of the words from the Speech Accent Archive, while treating L2 English speakers with different L1s as 'varieties'. This yields lists of words that are pronounced characteristically differently in comparison to native accents of English. We discuss English accent characteristics for French, Hungarian and Dutch, and compare the results to other sources of accent information. Knowing about these characteristic features of accents has useful applications in teaching L2 learners of English, since potentially difficult sounds or sound combinations can be identified and addressed based on the learner's native language.

1 Introduction

Dialectologists have taken advantage of computational techniques to study regional language variation, and developed specific measures for quantifying this

Jelke Bloem, Martijn Wieling & John Nerbonne. 2016. Automatically identifying characteristic features of non-native English accents. In Marie-Hélène Côté, Remco Knooihuizen & John Nerbonne (eds.), *The future of dialects*, 155–173. Berlin: Language Science Press. DOI:10.17169/langsci.b81.148

variation. This field of quantitative dialectology is known as dialectometry. Dialectometry research has led to a variety of methods for analyzing large numbers of dialectal features in systematic ways. In particular, aggregation of features made available new methods such as quantification of distances between dialects, and statistical analysis of differences that allowed generalization over the noise inherent in examining any single linguistic feature (Nerbonne 2009). However, dialectologists are still interested in examining single features as well. Typical characteristics of dialects, known as 'shibboleths', are quite salient and frequently discussed among both dialectologists and laymen. Prokić, Çöltekin & Nerbonne (2012) show that quantitative methods can provide insight into this phenomenon as well. They identify the most characteristic words for various Dutch dialects, providing statistical evidence due to the aggregation of data.

The methods that have been developed in dialectometry have not been widely applied to other domains of linguistics, but there are clear generalizations that can be made. Any time a set of language variants is studied, where the languages differ in a quantifiable way, dialectometry methods can potentially be applied. This is certainly the case in second language acquisition, where different language backgrounds lead to a lot of variety among learners. In the acquisition of a particular second language such as English, native Mandarin speakers will produce a different English than native German speakers. These kinds of differences can be studied with dialectometric methods.

In particular, researchers working on accent studies (i.e. Wells 1982; Waniek-Klimczak 2008) could benefit from the use of these methods. It has long been noted that foreign accents can be perceived negatively (Ryan 1983). As a consequence, pronunciation training is a part of second language teaching, in which the goal is to make the students' accents more native-like. Since it is quite difficult to achieve native-like proficiency in second language learning, it has long been acknowledged that learners do not need to learn how to speak perfectly, but that intelligibility is sufficient:

> The learner (…) would have presented to him certain carefully chosen features on which to concentrate, the rest of his pronunciation being left to no more than a general supervision (Abercrombie 1956: 93).

This suggestion has later been developed into the idea of a hierarchy of errors, i.e. pronunciation problems that require the most attention in pronunciation training. A summary of research in this direction is provided by van den Doel (2006: 7–15). He notes that such hierarchies "have been formulated partly on the basis of experimental research, but mainly as a result of impressionistic observational procedures". Obviously, they are also language-specific.

We are not aware of many studies that discuss error hierarchies of phonological errors, or characteristic feature rankings. One example of the use of error hierarchies in a more general sense can be found in Rifkin (1995). This analysis does not go to the level of phonological features, as it discusses grammatical errors and intonation errors. Gynan (1985) discusses phonological features and places them in an error hierarchy, but only on a general level. Based on data from Spanish learners of English and U.S. bilingual native speakers of English, he notes that comprehensibility of accents is related more to phonological than to morphosyntactic characteristics, but problems with morphosyntax are more salient to native speakers.

There are also studies that discuss characteristic pronunciation errors in English by speakers of a specific language. Gao (2005) studied a Chinese L2 student of English in a longitudinal study over 12 weeks, analyzing the errors and determining whether they arose from first-language interference or from being in an early stage of language acquisition. Potential errors were identified from earlier work on Chinese accents, a methodology that is strongly biased against the discovery of less stereotypical errors. The study finds that most errors arise from Chinese interference, though this may be partly due to the bias towards typical Chinese errors. The article also notes the need for research that studies a wider range of speakers.

Another line of work that assumes strong interference effects and makes comparisons to native speaker phonology is automatic accent classification. These methods are often also based on the assumption that the non-native speaker replaces unfamiliar sounds in the second language with sounds from their native language, e.g. by Angkititrakul & Hansen (2006).

One error hierarchy that explicitly includes phonological errors can be found in the thesis of van den Doel (2006). He carried out a large study where native English speakers were asked to detect and evaluate Dutch pronunciation errors, to provide more empirical evidence for attitudes towards specific pronunciation errors for this combination of languages. We will compare this error severity hierarchy approach with our characteristic feature ranking approach, and show that this measure of severity is not the same as measuring characteristic features by comparing results of the two approaches.

Schaden & Jekosch (2006) discuss an interesting data set that has applications in identifying characteristic pronunciation errors: the CrossTown corpus, which contains transcriptions of speakers of several European languages pronouncing place names from other European countries. In Schaden (2004), a rule-based system for accent generation was created from this data set. Rules that encode

typical pronunciation errors by speakers of one language in another language were derived manually in this study. Automatic identification of these errors would probably be possible from this data set, but does not appear to have been attempted.

Automatic identification of characteristic features of accents may provide additional empirical evidence for pronunciation difficulties. Since by definition native speakers rarely produce these features, they are likely to stand out. We propose that Prokić, Çöltekin & Nerbonne's (2012) method for detecting characteristics of dialects can be used for detecting characteristics of accents. Based on transcriptions of accented English speech from the Speech Accent Archive (SAA, Weinberger & Kunath 2011), we demonstrate how such characteristic features of accents can be identified. We quantify the most distinctive deviations from the standard English pronunciation for several languages of which native speakers are included in the archive. Note, however, that the method can be used for any language of which transcriptions from native speakers are available. We then compare the segments we identify to phonological features from published literature that are said to be typical of the English accent of that language.

To illustrate the method, we discuss the results for three languages: French, Hungarian and Dutch. First, however, we will explain the measure we use to determine the characteristic features.

2 Measure

Wieling & Nerbonne (2011) proposed two measures to identify characteristic features of dialects. The first measure is REPRESENTATIVENESS, which they defined as how frequently the feature occurred within the dialect area. A high representativeness indicates that the differences between pronunciations within the dialect area are small. The second measure is DISTINCTIVENESS, which they defined as how characteristic the feature is for the dialect. A high distinctiveness indicates that the differences between pronunciations within and outside the dialect area are large.

These measures are comparable to Labov, Ash & Boberg's (2006: p.43) isogloss measures: REPRESENTATIVENESS is identical to their measure of homogeneity, and DISTINCTIVENESS is similar but not identical to their consistency measure. The differences are discussed by Wieling, Upton & Thompson (2014). Furthermore, the representativeness measure is similar to RECALL and distinctiveness to PRECISION, as used in information retrieval.

Prokić, Çöltekin & Nerbonne (2012) showed that even a single dialect word can be used to characterize a dialect area using these measures. The measures

proposed by Wieling & Nerbonne (2011) were generalized by Prokić, Çöltekin & Nerbonne (2012) in order to apply them (numerically) to the word level, rather than at the level of the individual features. Given that we are interested in the word level, we follow Prokić et al.'s definition. A further advantage of focusing on the word level is that phonetic context is taken into account. Non-native speakers are likely to use phonological rules from their native language, which may depend on context.

Prokić, Çöltekin & Nerbonne (2012) define the measures from a dialectological perspective in terms of sites and groups — a site is a location where a dialect sample is observed, and a group is a dialect area. Since we are working with accent data, instead we will use the terms speakers and languages — a speaker is one person included in the Speech Accent Archive, and a language is a group of speakers with the same native language.

A very representative feature shows little variation among the English accents of native speakers of one language, and a very distinctive feature shows a large difference between those speakers and native speakers of English. More formally, we assume a native language l, consisting of $|l|$ speaker samples, among a larger group of languages G consisting of $|G|$ speaker samples. G includes the speakers s that speak l as well as the s speaking other languages. In this work, we limit G to only include native speakers of the language of interest l and of English, since we would like to see what features are characteristic compared to native English. However, including more languages in G is possible too.

We also assume a measure of between-speaker difference d, with respect to a given feature f. For representativeness, we then calculate a mean difference $\bar{d}$ with respect to f within the language under investigation:

$$\bar{d}_f^l = \frac{2}{|l|^2 - |l|} \sum_{s,s' \in l} d_f(s, s') \tag{9.1}$$

To quantify distinctiveness, we calculate a mean difference $\bar{d}$ with respect to f from the speech of native English speakers:

$$\bar{d}_f^{\neg l} = \frac{1}{|l|(|G| - |l|)} \sum_{s \in l, s' \notin l} d_f(s, s') \tag{9.2}$$

Characteristic features are considered to be those where the difference between $\bar{d}_f^{\neg l}$ and $\bar{d}_f^l$ is relatively large. Following Prokić, Çöltekin & Nerbonne (2012), we normalize these measures by calculating the difference between their z-scores rather than just the raw difference:

$$\frac{\bar{d}_f^{\neg l} - \bar{d}_f}{sd(d_f)} - \frac{\bar{d}_f^{l} - \bar{d}_f}{sd(d_f)} \tag{9.3}$$

This normalizes the difference scores for each feature separately.

This measure is implemented in the publicly available Gabmap web application for dialectology (Nerbonne et al. 2011), and this is the implementation we used to conduct this research.[1] In this application, languages l are represented as clusters of $|l|$ speaker samples. We manually defined these clusters using the native language metadata from the Speech Accent Archive, not applying any of the automatic clustering techniques available in Gabmap to avoid errors.

As for the measure of between-speaker difference d, we used the Gabmap function for finding the aggregated Levenshtein distance between two speakers' transcriptions, described by Nerbonne et al. (2011). This dialectometric method has also been applied to accent studies before. Wieling et al. (2014) found a correlation of $r = -0.81$ between human native-likeness judgments and the Levenshtein distance between native and non-native English speech.

We have applied this measure to transcriptions of the words from the Speech Accent Archive, each time comparing speakers of one particular language to native English speakers. After applying the formula above to the pronunciation distances, we identify lists of words that are characteristically pronounced differently by the non-native speakers, in comparison to native accents of English. To verify the measure and obtain more detail, we examined the top of these lists more closely. For the top five words, we looked at the most frequently occurring transcribed forms of the word in language l to see if they are indeed different from native English speech and if these differences might be called characteristic.

3 Material

Our transcriptions are a subset of transcriptions extracted from the Speech Accent Archive (SAA, Weinberger & Kunath 2011). The SAA has been expanded since we extracted the transcriptions, but we have used this older dataset because it has been segmented and manually checked. The SAA is available at http://accent.gmu.edu and contains a large collection of speech samples in English from people with various language backgrounds, including both native and

[1] Available at: www.gabmap.nl

non-native speakers of English. Each speaker reads the same paragraph containing 69 words in English:

> *Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.*

While reading out a paragraph may not be the most accurate representation of one's pronunciation ability, this method of elicitation makes sure that there is a set of comparable transcriptions for all speakers. Furthermore, this paragraph has been designed to include the most common phonemes of English, and should be able to serve as a standard for that reason.

To show how much information these transcriptions contain, we provide some example transcriptions from the SAA below. These are the first lines of the elicitation text, spoken by four speakers with different language backgrounds. Example 4 was spoken by a 42-year-old American male from Pittsburgh (english1 on the website). Example 5 is a female Hungarian speaker, who lived in both the UK and the USA for 1.5 years (hungarian1). Example 6 is a male Dutch speaker from the Netherlands, who only spent one month abroad in the UK (dutch1). Lastly, example 7 is a female speaker from France, who spent four months in the USA (french1). More information about the speakers is also available in the archive, but for this study we have not taken any of this metadata into account, except for the native language.

(4) [pʰl̥iːz kʰɑlˠ stɛlə æsk ɚ ɾə bɹĩŋ ðiːz θĩŋz wɪθ ɚ fɹʌ̃m ðə stɔɹ] (English)

(5) [plis kol stalʌ æsk hɜ tu brĩŋ d̪is t̪ĩŋz̩ wit̪ hər frɔ̃m d̪ə stɔr] (Hungarian)

(6) [pliːs kɔl stɛ̨la ask hɜ tu bɹĩŋ ʔðɪs ʔθĩŋs wɪθ hɜ fɹom d̪ə stoɜ] (Dutch)

(7) [pʰliz̩ kʰɔl stɛla æsk hɜɹ tu bɹĩŋ zɪs θɾːŋks wɪθ hɜɹ fɹʌ̃m d̪ə stɒɹ] (French)

Even from these single examples, we can already observe some typical foreign accent characteristics. The English speaker strongly reduces the word *her*, which the non-native speakers seem to be more conservative about. The English and French speakers aspirate their unvoiced plosives at the start of the first two words ([pʰ]), while the Hungarian and Dutch speakers do not, since their native languages lack aspirated stops. The open back unrounded vowel [ɑ] is not present in standard Dutch or Hungarian, and none of the speakers use it in the

example. We observe these speakers replacing it with more closed varieties of the vowel. We also see that not all stereotypical accent characteristics are always present: both the Dutch and French speakers correctly produced dental fricatives in the sequence *things with*, even though these languages do not include dental fricatives and non-native speakers are known to have trouble with this sound. Furthermore, French does not have aspirated stops just like Dutch and Hungarian, yet the French speaker still produced one (as noted above), while the Dutch and Hungarian speakers did not. Some speakers may be better at English pronunciation than others, and have learned to correctly use foreign sounds. We do see the French speaker substituting [z] for [ð] in *these*, and she aspirates the /p/ and /k/ in the first two words, which is unusual in French. She also devoices the final consonant of *things*, but not in *please*, showing that we not only find variation among native speakers of the same language, but that we also find it within the speech of a single speaker. We can observe one peculiar phenomenon in the Dutch transcription, the glottal stops before *these* and *things*. No other speakers of Dutch or English show this, and no such phoneme is apparent in the sound file, so it appears to be a transcription error. For these and other reasons, it is insufficient to examine the speech of a single speaker in discussing 'characteristic' accents. By aggregating, our method will provide stronger evidence of the characteristic features of accented speech.

4 Results

In this section, we will discuss the results of applying our method to French, Hungarian and Dutch accents. We limit ourselves to showing the top five most characteristic words according to the method, and their two most common transcribed forms. We have examined the French accents because many samples are available in the archive, Hungarian because it has some unusual phonological phenomena that span word boundaries and may be hard to detect, and Dutch, because we can compare our measure to the empirically established pronunciation error hierarchy of van den Doel (2006).

4.1 French

There are 34 speakers of French in the data set, providing us with a large sample of different forms of the words. Table 1 shows the five most characteristic words of the French speakers, ranked by their difference score (see previous section). This is calculated over all of the tokens in the elicitation paragraph. For words

Table 1: Characteristic words of French native speakers

Rank	Word	Score	Characteristic forms	Native forms
1	to	1.26	tu (20/34 : 11/181)	ɾə (0/34 : 112/181)
			tŭ (5/34 : 8/181)	
2	into	1.05	ĩntu (21/34 : 25/181)	ĩntə (1/34 : 56/181)
			ĩn̪tu (4/34 : 0/181)	ĩnɾə (0/34 : 29/181)
3	call	0.88	kɔl (14/34 : 12/181)	$k^h ɑl^ɣ$ (0/34 : 48/181)
			kɔːl (3/34 : 0/181)	$k^h ɔl^ɣ$ (1/34 : 13/181)
4	small	0.78	smɔl (22/34 : 33/181)	$smɑl^ɣ$ (1/34 : 59/181)
			smol (4/34 : 1/181)	smɔl (22/34 : 33/181)
5	can	0.50	kæ̃n (13/34 : 3/181)	kə̃n (1/34 : 82/181)
			kæn (4/34 : 0/181)	k^hə̃n (0/34 : 23/181)

that occur multiple times in the paragraph, we will refer to their tokens with an index, i.e. *the* [2] for the second instance of the word 'the' in the text. For each word, we also list the two most frequent forms used by French speakers, and the two most frequent forms used by native English speakers. If one form is used overwhelmingly more often than the other ones, we only list one. Behind each form, we list their frequency of occurrence in the following format: (French usage ratio : native usage ratio). For instance, for the first ranked item *to*, we can see that 20 out of 34 French speakers used the form [tu], while 11 out of 181 native English speakers used this form. It is highly characteristic of French. Native English speakers generally use the weakened form [ɾə], while the French speakers do not.

In French, unstressed vowels tend to be pronounced, and French speakers would be unlikely to produce the form [ɾə] anyway. The [ə] does exist in French, but it is phonetically realized only under special circumstances. A word-final schwa is usually elided, and only pronounced when the next word starts with a consonant. However, in the orthography this sound always appears as an <*e*>. A similar effect can be observed for *into* (ranked 2nd). English speakers use the form [ĩntə], used 56 times, as well as other forms ending in [ə], which are only used by one of the French speakers. The French language does not have vowel reduction to [ə] in word-final position, so it makes sense that French speakers would deviate from standard English here.

For the word *call* (rank 3), we mainly observe the use of [ɔ] as the vowel, while the majority of the native speakers uses [ɑ]. In French, the vowel [ɑ] is used, but it is in the process of merging with [a] (Walker 2001: 60–62). Perhaps for this reason, the French native speakers use [ɔ] in their English. The [ɔ] pronunciation can be observed in the speech of some native speakers as well. British Received Pronunciation (RP) speakers would use [ɔ] here, and this dialect is prestigious. Furthermore, the same difference can be observed in the Dutch and Hungarian data, though not as strongly. It may be the case that [ɔ] is taught to second language learners of English in this context, explaining the effect. The same phenomenon occurs in *small* (rank 4), where there are even some instances of [o] in the French-accented speech.

To continue, we can see that French native speakers do not aspirate the initial consonants of *call* or *can* (5th), for there are no aspirated consonants in Standard French (Walker 1984: p. 35). In the fifth word, *can*, we can also observe the usage of [ǽ] or [æ] instead of [ə] by the French speakers. [æ] is not a phoneme of standard French (Walker 2001), however, it is the vowel used in the full American English form of *can*. It is likely that the speakers have mostly acquired this English sound, but have not or not yet learned to reduce it, as the native speakers do.

Some properties of accents are considered to be effects of being in an early stage of learning regardless of the native language. However, it appears that many of the characteristic differences we found in French accents can be traced back to the phonology of Standard French.

4.2 Hungarian

Our discussion of the Hungarian accent data will refer to the English pronunciation teaching guide of Nádasdy (2006), which contains specific information on errors and substitution by Hungarian native speakers of English. Table 2 shows the most characteristic words of the Hungarian speakers. The top-ranked word *these* indeed shows two properties that seem to be typical of Hungarian accents and follow from the phonology of the language.

First, the dental fricatives [ð] and [θ] do not exist in Hungarian. The language has dental sounds and fricatives, but no dental fricatives, and using dental fricatives is considered to be a speech defect. Hungarian learners of English are said to often perceive these sounds as [f] and [v], but in production, the typical mistake is to replace [θ] with [s] and [ð] with [d] (Nádasdy 2006: p. 71). This is also what we observe in our data: the words *these* (rank 1 and 4) and *the* (rank 5) show [ð] being replaced by [d̪]. Second, we observe that a majority of the Hun-

Table 2: Characteristic words of Hungarian native speakers

Rank	Word	Score	Characteristic forms	Native forms
1	these [1]	2.06	d̪is (5/7 : 3/181)	ðiːz (0/7 : 35/181)
			d̪iz̪ (2/7 : 0/181)	ðiz (0/7 : 19/181)
2	please	1.70	plis (4/7 : 1/181)	pʰliːz (0/7 : 39/181)
			pʰliːs (2/7 : 5/181)	pʰliːz̪ (0/7 : 31/181)
3	big	1.69	bik (5/7 : 0/181)	bɪg(0/7 : 77/181)
			bɪk (1/7 : 1/181)	
4	these [2]	1.55	d̪is (4/7 : 1/181)	ðiz (0/7 : 59/181)
			d̪iz̪ (1/7 : 1/181)	ðiːz (0/7 : 38/181)
5	the [1]	1.52	d̪ə (6/7 : 3/181)	ðə (0/7 : 97/181)
			də (1/7 : 0/181)	n̪ə (0/7 : 64/181)

garian speakers devoices the [z] in *these*, something the English speakers do not do. This is likely to be an effect of Hungarian regressive (or anticipatory) assimilation. When two obstruents in Hungarian are pronounced in sequence, the first one assimilates to the second one — if the second obstruent is voiceless, the first obstruent will be voiceless, too. This can also occur across word boundaries, as long as there is no phonological gap. In the original text, both instances of *these* are followed by the word *things*, which the speakers pronounce with [t̪] (7 times), [t] (1 instance) or [θ] (6 times), which are all unvoiced obstruents. In Hungarian, regressive assimilation would devoice the [z] of *these* here, and this is also what happens in their English pronunciation.

The word *big* (rank 3) shows another clear example of regressive devoicing, but with [g] devoicing to [k]. The context in the elicitation paragraph is *big toy frog*, and the Hungarian speakers mostly pronounce [tɔɪ] as the English speakers do, with the only differences being in aspiration of the [t]. Since the [t] is unvoiced, the devoicing of the [g] in *big* is regressive devoicing. The word *please* (rank 2) shows the devoicing before the unvoiced [k] of *call*, but also a difference in aspiration. There are no aspirated stops in Hungarian (Petrova et al. 2006).

When looking at these characteristic features, one might wonder whether speakers always apply final devoicing in English. This is not the case, however. For example, the word-final [d] in the sequence *red bags* is voiced by all Hungarian speakers. These cases are not characteristic of the Hungarian accent, as English also has no strict final devoicing and English speakers use the [d] as

well. The score of the word *red* is only -0.48, the fourth lowest, showing less difference between Hungarian and English native speakers than among the Hungarian speakers. This indicates a quite similar pronunciation to native English.

In summary, our data show that regressive devoicing and the lack of dental fricatives are typical of Hungarian English accents compared to native English speakers' accents.

4.3 Dutch

While it is interesting to have quantitative evidence for characteristic features that can be linked to the phonetics of the native language, this does not tell us much about the ranking of the features. How do we know that the top five words really contain the most characteristic features? We are not aware of any other work that ranks phonetic or segment-based features of accents using a computational measure, but we may be able to find some evidence in perception studies. van den Doel (2006) conducted a large study on how Dutch accents are judged by British (Received Pronunciation) and American native speakers, which was aimed at finding salient pronunciation errors. In his study, he presented native English speakers 32 sentences, each containing a single pronunciation error considered to be typically Dutch, based on a survey. The pronunciation of the sentences was native, except for the error. Not all of the errors are phonemic (and therefore relevant to our study), but the ones that are, are considered by the authors to be representative of a more general phonological error. In the study, van den Doel (2006: 292) established hierarchies of errors consisting of five classes of severity, and separately for British English and American English. The most severe errors according to both groups are stress errors, which are not relevant to our study. We will discuss the most severe phonemic errors mentioned in the study, reproduced in Table 3, and compare them to our most characteristic features of Dutch accents, the top five of which are listed in Table 4.

For the American English data, two phonemic errors were classified in the most severe error class van den Doel (2006): the use of the uvular trill [ʀ], and 'fortis/lenis neutralization' (similar to devoicing, replacing [v] with [f], [d] with [t]). The first error is not observed in our top five. The topmost word where an [ʀ] might be found is *for* [2] at rank 19. However the Dutch speakers either use [ɹ] or no final consonant at all, and this is similar to what the native speakers do. In fact, in all of the words spoken by the 16 Dutch native speakers in the SAA, no instances of [ʀ] occur. The error may be severe and distinctive, but not representative, and therefore not characteristic. Even in native Dutch, [ʀ] is only used in the south, and throughout the Dutch language area, five main categories of *r* are

Table 3: Dutch hierarchy of error including only errors of severity > 2.2, adapted from van den Doel (2006)

Severity	Received Pronunciation	General American
> 3.5	Stress errors	Stress and stress-related errors
		Fortis/lenis neutralization
		Use of uvular-**r**
$2.2 - 3.5$	Stress-related errors	Most substitutions of /θ,ð/ by /t,d/
	Fortis/lenis neutralization	Glottalisation of final /d/
	Use of uvular-**r**	Epenthetic [ə] in /lm/
	Some substitutions of /θ,ð/ by /t,d/	/v ∼ w/ confusion
	Glottalisation of final /d/	/æ ∼ e/ confusion
	Epenthetic [ə] in /lm/	Inappropriate post-vocalic **r**
	/v ∼ w/ confusion	
	Confusion of /æ ∼ e, ʌ ∼ ɒ, ʊ ∼ u:/	
	Unaspirated [t]	

in use, with further subdivisions possible (Sebregts 2015). This illustrates the fact that the perception experiment of van den Doel (2006) identifies perceptually salient errors, which do not necessarily have to be characteristic. To be ranked highly by our measure, a feature has to be used by many non-native speakers of the language under consideration. Nevertheless, the second error, fortis/lenis neutralization, does occur. While *of* (rank 5) is more commonly pronounced with a final [v] by the native English speakers, all Dutch speakers use [f]. In our ranking, we also observe other forms of devoicing at the end of the words *slabs* (rank 3), *bags* (rank 4) and *big* (rank 1), though *big* is more likely to stem from the fact that most Dutch speakers do not use [g]. These phenomena were not included in the study of van den Doel (2006) at all, so it is unclear whether these forms are perceived as severe errors.

Table 4: Characteristic words of Dutch native speakers

Rank	Word	Score	Characteristic forms	Native forms
1	big	1.92	bɪk (13/16 : 1/181)	bɪg (0/16 : 77/181) bɪgˀ (0/16 : 41/181)
2	to	1.22	tu (10/16 : 11/181) tə (3/16 : 21/181)	ɾə (0/16 : 112/181)
3	slabs	1.12	slæps (5/16 : 0/181) slæb̥s (3/16 : 1/181)	slæbz (1/16 : 66/181) slæːbz (0/16 : 38/181)
4	bags	1.08	bæks (4/16 : 0/181) bæːg̊s (3/16 : 2/181)	bægz (1/16 : 39/181) bæːgz (0/16 : 33/181)
5	of [1]	1.06	ɔf (7/16 : 7/181) əf (7/16 : 46/181	əv (0/16 : 58/181) əf (7/16 : 46/181)

Out of the remaining errors listed in the second-most severe class for both British and American English, all but one appear to be relatively uncharacteristic for Dutch non-native speakers of English. /v/-/w/ confusion is listed as severe, and might be expected because the Dutch /w/ is usually often pronounced [ʋ]. It almost never occurs in the data. For each word containing a *w*, all but one or two Dutch speakers use [w]. In fact, both instances of *we* in the elicitation paragraph are the two lowest ranked words using our difference scores. Another such confusion, /æ/-/e/ confusion, might be expected in *slabs* (rank 3 in Table 4). This confusion may arise because Dutch does not normally use [æ]. However, no Dutch speakers pronounce *slabs* with an [e], though [a] and [ɛ] are each used by one speaker. The word mainly ranks highly because of devoicing in the final consonant cluster. Another error considered severe, though it can only occur in intermediate stages of learning, is glottalization of final /d/. While Dutch does not have any final glottalization, the hypothesis is that because English has glottalization of final /t/, Dutch speakers may generalize it. There is only one final /d/ in our elicitation paragraph, in *red*, and the phenomenon does not occur there. The word has a score of only -0.05, indicating a very similar distribution of forms as among the native speakers, with the exception of two speakers who did fortis/lenis neutralization. The last uncommon error is the insertion of an epenthetic schwa in [lm] clusters. There are no such consonant clusters in the elicitation paragraph, and there are not many words that end in *lm* in English.

One notable characteristic that van den Doel (2006) classifies in the second-most severe category, is the replacement of dental fricative with other sounds

(most likely [t] and [d] in Dutch). The highest ranked word with a dental fricative in Dutch accents is *the*[2] at rank 17. The reason is that various different replacements of the [ð] are used by the Dutch speakers. The expected phoneme [d] was used seven times, and four times [d̪] was used, the dental variety which is used in Flemish Dutch. Only two out of these four speakers were actually Flemish, so it may be used as a closer approximation of a dental fricative when learning English. Three more speakers correctly used [ð]. Since there is so much variation, there is no form that is particularly representative of Dutch accents, and the feature is not judged to be characteristic as a result. However, if we rank the Dutch accent features only by distinctiveness, two instances of *the* are ranked second and third. Consequently, this approach may be use to detect errors which show great variability by the non-native speakers.

5 Discussion

In this paper, we have demonstrated the use of dialectometric techniques to study English accents. We hope to have shown that methods from dialectology can be applied to other domains of linguistics in which there is language variation. We have used a quantitative measure to identify characteristic features of the accents of several languages. By aggregating over the transcriptions of multiple speakers from the Speech Accent Archive, we obtain stronger evidence for these features than one would obtain from the analysis of single transcriptions. We verified the resulting feature rankings by comparing them to three other sources of information relevant to accents: phonologies of the native language (Walker 2001), pronunciation teaching literature (Nádasdy 2006), and an empirical error perception study (van den Doel 2006). From the phonology and pronunciation literature, we learned that most of the characteristic features that we found are indeed a direct effect of interference from the native language, as opposed to some intermediate stage of learning. Furthermore, our method provides quantitative evidence for these observations, something we were not able to find in other work. It also yields a ranking of the words that phonological features occur in, providing more detail than was previously possible.

In the comparison to the perception study, we observed that our measure of characteristic features only somewhat overlaps with the perceived severity of speech errors. In particular, uncommon differences may be severe, but not characteristic due to their rarity. Difficult phonemes that are substituted in various ways by different speakers of an accent, are not deemed characteristic by our method. To identify these errors, the distinctiveness measure can be used.

The identification of characteristic features of accents can provide an additional source of information for teachers of English, since the measure favours features that are widespread, as opposed to some of the more stereotypical errors described by van den Doel (2006). They also differ from these stereotypical errors, indicating that our method may find errors that are not typically considered by teachers. By obtaining these characteristics in an empirical, objective and reproducible way, existing insights on L1-specific pronunciation errors can be validated against a dataset of transcriptions. Our method can also identify characteristic features of non-native speakers in other languages, as long as transcriptions of the SAA elicitation paragraph are available. This information can be applied in teaching L2 learners of English. Potentially difficult sounds or sound combinations can be identified and addressed based on the learner's native language.

One limitation of the method is that we still require a manual step to find phonological features in the transcribed forms of the words. In future work, perhaps this method can be combined with identifying characteristic sound correspondences (Wieling & Nerbonne 2011). An obvious continuation of this line of work is to apply this method to English accents of other languages. Finally, we suggest that dialectometric methods could be applied to the study of accents more often, since the two fields have many common characteristics.

Acknowledgements

We would like to thank Anna Mészáros for suggestions regarding the Hungarian data, and the anonymous reviewers for their helpful comments.

References

Abercrombie, David. 1956. *Problems and principles: Studies in the teaching of English as a second language*. London, New York, Toronto: Longmans, Green.

Angkititrakul, Pongtep & John HL Hansen. 2006. Advances in phone-based modeling for automatic accent classification. *Audio, Speech, and Language Processing, IEEE Transactions on* 14(2). 634–646.

Gao, Lili. 2005. Pronunciation difficulties analysis: A case study using native language linguistic background to understand a Chinese English learner's pronunciation problems. *Celea Journal* 28(2). 76–84.

Gynan, Shaw Nicholas. 1985. Comprehension, irritation and error hierarchies. *Hispania* 68. 160–165.

Labov, William, Sharon Ash & Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology and sound change*. Berlin, New York: Mouton de Gruyter.

Nerbonne, John. 2009. Data-driven dialectology. *Language and Linguistics Compass* 3(1). 175–198.

Nerbonne, John, Rinke Colen, Charlotte Gooskens, Peter Kleiweg & Therese Leinonen. 2011. Gabmap – a web application for dialectology. *Dialectologia* Special Issue II. 65–89.

Nádasdy, Ádám. 2006. *Background to English pronunciation*. Budapest: Nemzeti Tankönyvkiadó.

Petrova, Olga, Rosemary Plapp, Catherine Ringen & Szilárd Szentgyörgyi. 2006. Voice and aspiration: Evidence from Russian, Hungarian, German, Swedish, and Turkish. *The Linguistic Review* 23(1). 1–35.

Prokić, Jelena, Çağri Çöltekin & John Nerbonne. 2012. Detecting shibboleths. In *Proceedings of the EACL 2012 Joint Workshop of LINGVIS & UNCLH*, 72–80. Association for Computational Linguistics.

Rifkin, Benjamin. 1995. Error gravity in learners' spoken Russian: A preliminary study. *The Modern Language Journal* 79(4). 477–490.

Ryan, Ellen Bouchard. 1983. Social psychological mechanisms underlying native speaker evaluations of non-native speech. *Studies in Second Language Acquisition* 5(02). 148–159.

Schaden, Stefan. 2004. CrossTowns: Automatically generated phonetic lexicons of cross-lingual pronunciation variants of european city names. In *Proceedings of LREC 2004*, 1395–1398.

Schaden, Stefan & Ute Jekosch. 2006. Casselberveetovallarga and other unpronounceable places: The CrossTowns corpus. In *Proceedings of LREC 2006*, 993–998.

Sebregts, Koen. 2015. *The sociophonetics and phonology of Dutch r*. Utrecht: LOT.

van den Doel, Rias. 2006. *An evaluation of native-speaker judgements of foreign-accented British and American English*. Utrecht: LOT.

Walker, Douglas C. 1984. *The pronunciation of Canadian French*. Ottawa: University of Ottawa Press.

Walker, Douglas C. 2001. *French sound structure*. Vol. 1. Calgary: University of Calgary Press.

Waniek-Klimczak, Ewa. 2008. *Issues in accents of English*. Vol. 2. Newcastle-upon-Tyne: Cambridge Scholars Pub.

Weinberger, Steven H & Stephen A Kunath. 2011. The speech accent archive: Towards a typology of English accents. *Language and Computers* 73(1). 265–281.

Wells, John C. 1982. *Accents of English*. Vol. 1. Cambridge: Cambridge University Press.

Wieling, Martijn & John Nerbonne. 2011. Bipartite spectral graph partitioning for clustering dialect varieties and detecting their linguistic features. *Computer Speech & Language* 25(3). 700–715.

Wieling, Martijn, Clive Upton & Ann Thompson. 2014. Analyzing the BBC Voices data: Contemporary English dialect areas and their characteristic lexical variants. *Literary and Linguistic Computing* 29(1). 107–117.

Wieling, Martijn, Jelke Bloem, Kaitlin Mignella, Mona Timmermeister & John Nerbonne. 2014. Measuring foreign accent strength in English. Validating Levenshtein distance as a measure. *Language Dynamics and Change* 4(2). 253–269.