Abstract

In this paper I argue that textual data from historical periods can be evidence of indexicality (Silverstein 2003) and enregisterment (Agha 2003), and that repertoires of enregistered features can be identified in historical contexts. Using a corpus of nineteenth-century Yorkshire dialect material, I discuss a repertoire of features that was enregistered as “Yorkshire” to nineteenth century audiences. I do this by comparing data from a corpus of modern Yorkshire dialect material and an online survey of current speakers’ perceptions of Yorkshire dialect. I suggest that similar patterns as those seen in the modern Yorkshire dialect data can be observed in analogous data from the nineteenth century. I argue that we can therefore make ‘use of the present to explain the past’ following Labov (1977: 226) and infer that features of Yorkshire dialect which occur frequently in the historical corpus and metapragmatic discourse were enregistered in the nineteenth century.

Keywords
indexicality, enregisterment, Yorkshire, dialect, corpus

EL “ENREGISTERMENT” EN CONTEXTOS HISTÓRICOS: EL DIALECTO DE YORKSHIRE EN EL SIGLO XIX

Resumen

En este artículo se argumenta que los datos textuales de períodos históricos pueden evidenciar “indexicalidad” (según Silverstein 2003) y “enregisterment” (según Agha 2003), y que los repertorios de rasgos registrados pueden identificarse en contextos históricos. Usando un corpus del dialecto de Yorkshire
del siglo XIX, se estudian determinadas características que fueron consideradas de "Yorkshire" por los hablantes del siglo XIX. Este objetivo se lleva a cabo mediante la comparación de los datos con un corpus del dialecto moderno de Yorkshire y con una encuesta en línea sobre las percepciones actuales de los hablantes de este dialecto. Se sugiere que patrones similares a los observados en el dialecto moderno de Yorkshire pueden observarse en datos análogos del siglo XIX. Se sostiene, pues, que se puede "utilizar el presente para explicar el pasado", siguiendo a Labov (1977: 226), e inferir que las características del dialecto de Yorkshire que aparecen con frecuencia en el corpus histórico y en los discursos metapragmáticos se registraron ya en el siglo XIX.

Palabras clave
“indexicalidad”, “enregisterment”, Yorkshire, dialecto, corpus

1. Introduction

Enregisterment was defined by Asif Agha as “processes through which a linguistic repertoire becomes differentiable within a language as a socially recognized register of forms” (Agha 2003: 231). Johnstone et al. develop this notion further, stating that a feature has been enregistered when it has “become associated with a style of speech and can be used to create a context for that style” (Johnstone et al. 2006: 82). Agha goes on to state that the cultural value associated with certain enregistered forms is “a precipitate of sociohistorically locatable practices, including discursive practices, which imbue cultural forms with recognizable sign-values and bring these values into circulation along identifiable trajectories in social space” (Agha 2003: 232), suggesting that the discursive practices involved in enregisterment may be identifiable in historical periods as they are now. I argue that it is therefore possible to study enregisterment in historical contexts.

In this paper I examine textual evidence concerning the Yorkshire dialect from the nineteenth century and discuss whether textual representations of dialect features and historical metapragmatic discourse can highlight these features’ enregisterment. I then go on to consider patterns observable in dialect material for modern Yorkshire dialect and how this, combined with metapragmatic data from an online survey of current speakers, can inform our interpretations of the historical data. Ultimately, I address the issue of whether historical textual data is sufficient to discuss enregisterment in historical contexts.
1.1 Enregisterment in historical contexts

Beal states in her discussion of the enregisterment of two varieties of Northern English, “Geordie” in Newcastle and “Sheffieldish,” that the repertoire of Tyneside (Newcastle) English was enregistered “early enough for it to be used in performative contexts in the nineteenth-century music halls” (Beal 2009: 140). This suggests that there may be specific evidence to directly highlight the nineteenth-century enregisterment of this variety.

Furthermore, Aaron’s study of the historical development of linguistic stereotypes in literary media also suggests evidence for the observation of enregisterment in historical periods. Using the CORDE corpus of written Spanish, she considers literary representations of non-standard Spanish dialectal forms over a period of almost 1,000 years. She states that a rise in frequency non-standard dialect representations in literary media is “indicative of these forms’ arrival, or literary/representational ‘re-birth’ as linguistic stereotypes” (Aaron 2009: 474). She goes on to state that there is a diachronic pattern to the development of linguistic stereotypes in literature over time, characterised by “a slow decline in frequency, nearly to extinction, followed by a sharp and exaggerated rise” (Aaron 2009: 476). She concludes that by considering both the quantitative frequencies of certain features and qualitative metadata from historical periods, we may be able to arrive at “tentative conclusions about the social meaning of other variants in societies long gone” (Aaron 2009: 492). Aaron does not specifically discuss the concepts of indexicality (Silverstein 2003) or enregisterment, though; she discusses linguistic stereotypes in relation to Labov’s paradigm for “indicators”, “markers” and “stereotypes”, where stereotypes are defined as “socially marked forms, prominently labelled by society” (Labov 1972: 314). However, if we follow Johnstone et al. and equate Labov’s notion of a linguistic stereotype with third-order indexicality (Johnstone et al. 2006: 82-3), then we can also interpret Aaron’s conclusions about social meaning from textual representations and historical metadata as evidence for these variants’ enregisterment.

Finally, Ruano-García discusses the enregisterment of the northern dialect of Early Modern English as represented in literary texts. He argues that there is a set repertoire of features used in literary representations of Northern English, highlighted by quantitative...
corpus analysis, and concludes that “EModE ballads and plays show enregisterment of some linguistic features”, and that “a specific set of forms is fairly consistently used in both ballads and drama” (Ruano-García 2012: 381) when representing Northern English. He goes on to state that the wide distribution of these texts in the Early Modern Period led to these Northern forms being brought into contact with non-Northern speakers, which created “a collective linguistic idea about the dialect itself” (Ruano-García 2012: 381); however, unlike Aaron above, Ruano-García does not discuss any form of qualitative metadata for Northern English in the Early Modern Period. I discuss the possibility of studying enregisterment using both quantitative and qualitative data from historical textual sources further and in more detail below.

1.2 Historical context: nineteenth-century Yorkshire

The nineteenth century forms the historical context for this paper for several reasons. Firstly, the Industrial Revolution gave rise to great cultural and social change in England. The increased ease of social mobility and the emergence of the middle class led to “white-collar and service based jobs” which “demanded a veneer of gentility” (Beal 2004: 179). This “genteel” veneer was in part represented by “correct” pronunciation, and resulted in a demand for guides detailing “correct” English usage. Jones notes that certain shibboleths and “mis-pronunciations” were focused on as being “vulgar and low-class” (Jones 2006: 286), stating that these usage guides were not aimed at the lowest classes of society, but instead were aimed at the “newly socially aspirant” (Jones 2006: 287), who wanted to distinguish themselves from the “vulgar” and “low-class” strata of society. The effect of this linguistic awareness on the part of the “newly socially aspirant”, however, “was the creation of what Labov has termed the ‘linguistic insecurity’ typically associated with the middle class” (Beal 2004: 94).

Secondly, the Industrial Revolution was also responsible for several technological advances, particularly that of the railways, which made geographical mobility much easier and faster. Until mid-century, the canal system was the easiest and one of the more prosperous means of transporting cargo around the country. By the 1830s and 1840s, though, the railways began to expand, creating a cheaper, faster, and far more efficient
means of transport. It soon also became apparent that rail travel could enable people to move much more freely around the country, and “Thomas Cook was the first man – in 1859 – to see the country’s potential for ‘tourism’” (Robbins 1989: 25). One result of this geographical mobility was greater linguistic awareness and, as Wales states, dialect “found a new medium printed on the postcard home” (Wales 2006: 137).

Nineteenth-century awareness of regional dialects was also directly brought about by industrialisation; Wales goes on to state that “most Northern cities grew on the strength of the incoming populations from their rural hinterlands” (Wales 2006: 115-116). Population movement from rural areas towards newly-industrialised urban centres created a fear that (particularly rural) dialects would “die out”, as discussed by Milroy, who states that “strong interest in English ... dialects developed in the nineteenth century” (Milroy 2002: 14). This was particularly the case with regards to the Yorkshire dialect; we can see evidence of contemporary observations to this effect in the form of those made by antiquarian Joseph Hunter, who stated that “more attention has been paid to the verbal peculiarities of Yorkshire than of any other county” (Hunter 1829: xx).

We can therefore observe in the nineteenth century two of the key causal factors which can allow processes of enregisterment to occur: geographical and social mobility. In order for features to shift orders of indexicality and become enregistered, Johnstone et al. argue that these kinds of mobility can be contributing factors to the process. They state that social mobility can shift speakers’ awareness of features from first to second-order indexicality as “the choice among variants could, for some people, be invested with second-order indexical meaning such as class or correctness” (Johnstone et al. 2006: 89). For third-order indexicality, geographical mobility is required. When this happens, according to Johnstone et al., wider awareness of second-order variation can occur in two ways: firstly, as speakers from a particular community leave and go elsewhere, their linguistic variants are linked to their geographical region; secondly, when speakers move into a community from out of the region, local speech forms are noticed and again linked to place (Johnstone et al. 2006: 93-94).

The focus of this paper is on the potential for observation of enregisterment of nineteenth-century Yorkshire dialect (see also Cooper 2013, 2014), as the industrialisation of much of Yorkshire, particularly the West Riding, led to unprecedented population rises
as people moved from rural areas to urban centres (see Wright 1986; Feinstein 1981). In addition, the advance and spread of the railway network led not only to geographical mobility generally throughout the country, but to the prosperity of certain industrialised areas like Hull and York, both of which greatly benefitted from the railways in terms of trade and employment respectively (see Gillett & MacMahon 1980; Feinstein 1981); whereas other areas of the county became prosperous as holiday and leisure destinations, like Scarborough (Singleton 1970: 52). These developments would ultimately lead to increased awareness of the features of the “Yorkshire” dialect.

1.3 Nineteenth-century Yorkshire dialect data

The nineteenth century saw a tremendous output of dialect representation, particularly in Northern England. Beal notes that the “growth of urban population in towns and cities such as Newcastle, Manchester, and indeed Halifax, led to the creation of a market for popular forms of literature in dialect” (Beal 2004: 204). Indeed, representations of Yorkshire dialect in the nineteenth century can be observed in the form of both dialect literature and literary dialect (Shorrocks 1996: 386), where the former includes works written entirely in dialect (for instance, poems, ballads, songs, dialogues); the latter is dialect represented in novels and plays (for instance, the dialogue of the character Browdie in Dickens’ *Nicholas Nickleby*). The distinction between these two forms of dialect representation lies in the general tendency for dialect literature to be written by local writers for local audiences, as discussed by Leith & Graddol, who state that dialect literature was “both printed and sold by local publishers. Many of the dialect writers were workers and they were often self-educated in the new textile factories of Lancashire and Yorkshire” (Leith & Graddol 2002: 162). It is likely that much of this work was aimed at local audiences and there was considerable demand for these kinds of texts. This is as opposed to literary dialect, which was often written by non-local writers and intended for a wide audience; for instance, Charles Dickens was not native to Yorkshire, yet represents Yorkshire dialect in *Nicholas Nickleby* (1838). Dickens’ works were hugely popular in the nineteenth century, and were read by ‘a very large and diverse audience’ (Pykett 2002: 3), making him “a world-famous author” (Patten 2001: 16) in his own
lifetime. We can therefore compare representations of Yorkshire dialect written by local speakers for local audiences with those written by non-local speakers for international audiences. This is important for studying the nineteenth-century enregisterment of the Yorkshire dialect as these sources can highlight which features are salient in representing Yorkshire at a local and international level.

In addition, we can also observe nineteenth-century metapragmatic discourse on Yorkshire dialect. This is the closest analogue to explicit “talk about talk” described by Johnstone et al. (2006: 93) in the absence of any living speakers of Yorkshire dialect from the nineteenth century. This metapragmatic dialect ‘commentary’ material comes from various sources including: introductory material for dialect dictionaries, essays about particular dialects, travel writing, articles from popular magazines, books written about particular dialects, dialect glossaries, and dialect grammars. Many different aspects of the Yorkshire dialect were commented upon by contemporary observers; we can see comments such as: “asking where we should look we say weer” (Piper 1824: 10 – italics in original), on the Yorkshire pronunciation of the SQUARE diphthong as the NEAR diphthong /ər/; and “The absence of þ or th in the definite article is remarkable” (Addy 1888: xviii), highlighting the use of definite article reduction [DAR] in the Yorkshire dialect. Some commentators are somewhat more general in their discussion of the dialect; for instance, Fisk, an American Methodist preacher, records in his travel writing of his time spent in Yorkshire: “Their prepositions and conjunctions are mixed up and interchanged for each other in such grotesque order, and their vowels are sounded so queerly, that every sentence is amusing” (Fisk 1838: 669). This explicit metacommentary on the features of the Yorkshire dialect serve to link language features with place, succinctly summarised by contemporary commentator Morris, who cites the phrase “it takes a Yorkshireman to talk Yorkshire” (Morris 1892: 44) in his discussion of the “intelligibility” of Yorkshire dialect to “foreigners” (i.e. anyone who is not native to Yorkshire).

1.4 Corpus of nineteenth-century Yorkshire dialect material

The historical data for this paper comes from a corpus I constructed of nineteenth-century dialect material concerning Yorkshire dialect. This corpus is in two main sections:
(i) qualitative metacommentary featured in 18 texts from the types of sources listed above; and (ii) samples of dialect representation from 27 dialect literature texts (26,376 words) and 20 literary dialect texts (18,229 words) which amount to 44,605 words in total for quantitative analysis.

Analysis of the qualitative data highlighted that several Yorkshire dialect features were consistently discussed in the ‘commentary’ material. These included:

(i) Definite Article Reduction
(ii) Alternate diphthongs (coit ‘coat’; weer ‘where’)
(iii) Alternate vowels (<i> for <u> sich ‘such’)
(iv) /l/-vocalisation (oud/owd ‘old’)
(v) Lexical items (owt ‘anything’; nowt ‘nothing’; summat ‘something’; shoo/hoo ‘she’; bairn ‘child’; gan ‘go’; sen ‘self’; mun ‘must’; nobbut ‘only’)

Quantitative analysis of the dialect literature and literary dialect also highlighted several features that were similarly consistently and frequently employed in representations of Yorkshire dialect. These included several of the features that were consistently discussed in the qualitative material listed above:

(i) Definite Article Reduction
(ii) Alternate diphthongs (ageean ‘again’; reight ‘right’)
(iii) Alternate vowels (<i> for <u> sich ‘such’; <o> for <a> mony ‘many’)
(iv) l-vocalisation (oud/owd ‘old’)
(v) Lexical items (owt ‘anything’; nowt ‘nothing’; summat ‘something’; shoo/hoo ‘she’; bairn ‘child’; gan ‘go’; sen ‘self’; mun ‘must’; nobbut ‘only’)

In addition, several of the features which were consistent in the quantitative data appeared in similar proportions in both the dialect literature and the literary dialect, as shown in Figure 1.

We can see in Figure 1 that the features: mun ‘must’; <ea> in theare ‘there’; nowt ‘nothing’; gan/gang ‘go’; <th’>; <oa> in goa ‘go’; sen ‘self’; bairns ‘children’; <t’>; and <o> for <a> in onny ‘any’ all appear in similar proportions in both types of dialect
representation, as they occur within a percentage proportion range of around 40:60 to 70:30 dialect literature [DL] to literary dialect [LD]. If we are to continue to assume that DL is predominantly produced for and received by a local Yorkshire audience, and that LD is aimed at a wider, potentially international audience, then the fact that the features listed above appear in almost equal proportions of DL:LD suggests that those features are almost equally salient in representing Yorkshire dialect to a local audience as to a non-local one.

The analysis of the historical data led me to the following hypothesis. Features which: (i) are consistently discussed in the qualitative material; (ii) are quantitatively numerous and consistently frequent in dialect representations in DL and LD; and (iii) appear in similar quantities in DL and LD, therefore suggesting similar salience to local versus non-local speakers; may have been enregistered to nineteenth-century audiences (see also Cooper 2013, 2014).

![Figure 1. Mean averages of commonly-occurring dialect features in dialect literature versus literary dialect across nineteenth-century corpus.](image-url)
In order to test this hypothesis, and in lieu of any surviving speakers of nineteenth-century Yorkshire dialect, I constructed an analogous corpus of modern dialect material and compared the results with an online survey of modern speakers.

2. Corpus of modern Yorkshire dialect material

The modern corpus was designed to be as analogous to the nineteenth-century material as possible. As a result, when selecting qualitative data, sources in media that are not traditionally published such as web-based sources, were excluded; for instance internet pages such as a wikiHow entry on “How to speak with a Yorkshire accent” (http://www.wikihow.com/Speak-With-a-Yorkshire-Accent), or Youtube videos which discuss the ‘Yorkshire’ accent and dialect such as “How To: Be a Proper Yorkshire Pud” (http://www.youtube.com/watch?v=6V54g0314UA). The modern qualitative data therefore came from 7 texts, predominantly in the form of prefatory material from dialect glossaries and dictionaries. The quantitative data was comprised of a smaller corpus of 5,000 words of DL and 5,000 words of LD; this corpus was not as extensive as that for the historical data, as the modern corpus data was to be augmented with the results from the online survey. There was also an apparent lack of available modern quantitative data, particularly DL, when compared with the nineteenth century.

Analysis of the modern corpus showed that, just as with the historical data, there were several features which were both consistently discussed in the qualitative material and frequently used in the quantitative data. These included:

(i) Definite Article Reduction
(ii) /h/-dropping
(iii) Representations of Yorkshire phonological features (*allus/alis* ‘always’; *o’er/ower/ovver* ‘over’)
(iv) Alternate diphthongs (*rooad* ‘road’; *weer* ‘where’)
(v) Lexical items (*owt* ‘anything’; *nowt* ‘nothing’; *summat* ‘something’; *sen* ‘self’; *thee/tha* ‘you’)

©Universitat de Barcelona
These features also fell within similar proportions of DL:LD in a similar manner to the features that were frequent and consistent in the historical corpus.

3. Online Survey

The online survey was created using the online web service www.kwiksurveys.com. Respondents were asked to provide their age, gender, and whether or not they were from Yorkshire; a box was provided for a more specific response to this question. If respondents were from Yorkshire, they were asked to state where within the county they were from, in a format such as “Sheffield, South Yorkshire”. Respondents not from Yorkshire were also asked for a specific location in the similar form of “town/city, county”. This created three groups of respondents: “Yorkshire”, “non-Yorkshire”, and “International”; the latter group were respondents from outside of the mainland UK. Respondents were also grouped by age in the brackets 18-29; 30-39; 40-49; 50-59; 60 and over. The survey was split into two parts; in part 1, respondents were asked to provide, if possible, up to ten features they felt were representative of the Yorkshire dialect; these could be words, pronunciations, phrases, etc. They could progress no further in the survey until this was complete; the second part of the survey was unseen to the respondents at this stage and on a separate page. Part 2 of the survey was a multiple-choice exercise where respondents were asked to rate 23 features from both corpora according to several criteria which highlighted the strength of the association of those features with Yorkshire.

The survey ran for 1 month; in that time there were 410 respondents. Of those 410 respondents, the group with the smallest number was female Yorkshire speakers aged 60 and over. Only 7 respondents filled in the survey from this group; of those 7, only 6 actually completed the survey. As a result, 12 respondents were chosen from each age group for analysis (6 male and 6 female), giving a total of 120 respondents analysed from the mainland UK. There were an additional 21 respondents who were from overseas; these results were considered as a separate group due to the smaller number of
respondents, and relative inconsistency of locations represented in the ‘International’ group.

- Survey results

The survey results highlighted that there was a considerable amount of consistency in the "Yorkshire" features respondents provided; similar consistency was also displayed in the association of certain features from the corpora with Yorkshire in the multiple choice exercise. Table 1 shows the features which, on average, were consistently and frequently listed by all groups of respondents in decreasing order of frequency.

The following features were those which, on average, respondents most strongly associated with Yorkshire in the multiple choice exercise, in decreasing order of association strength:

(i) Definite Article Reduction
(ii) *Nowt* ‘nothing’
(iii) *Owt* ‘anything’
(iv) *Summat* ‘something’
(v) */h/-doloping
(vi) *Reight* ‘really/right’
(vii) *Sen* ‘self’
(viii) *Thee/tha* ‘you’
(ix) *Ower* ‘over’
(x) *Allus* ‘always’

<table>
<thead>
<tr>
<th>“Yorkshire” Features Provided</th>
<th>Average Percentage of all Respondents</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Thee/tha</em> ‘you’</td>
<td>54</td>
</tr>
<tr>
<td>Definite Article Reduction</td>
<td>45</td>
</tr>
<tr>
<td><em>Reight</em> ‘really’</td>
<td>33</td>
</tr>
<tr>
<td><em>Nowt</em> ‘nothing’</td>
<td>33</td>
</tr>
<tr>
<td>Terms of Endearment (e.g. ‘love’, ‘duck’)</td>
<td>24</td>
</tr>
<tr>
<td>*/h/-dropping</td>
<td>24</td>
</tr>
</tbody>
</table>
When we compare the results from the 2 parts of the online survey, we can see that there are seven features which appear in both of the above lists. These features are:

(i) Definite Article Reduction  
(ii) *Nowt* ‘nothing’  
(iii) *Owt* ‘anything’  
(iv) */h/-dropping  
(v) *Reight* ‘really/right’  
(vi) *Sen* ‘self’  
(vii) *Thee/tha* ‘you’

The common features listed above also displayed the following tendencies in the modern corpus data, in that they all:

1. featured direct commentary in 40% or more of the qualitative texts  
2. featured tokens in 50% or more of the quantitative corpus texts  
3. occurred within a percentage proportion range of 40:60 to 90:10 DL:LD

This highlights a correlation between the textual data from the corpus and the features listed as “Yorkshire” by respondents to the online survey.

4. Conclusions

By comparing the survey data with the modern corpus data, we can see that there is a strong correlation between frequently and consistently-occurring features in the modern corpus and the survey respondents’ perceptions of salient “Yorkshire” features.
This suggests that the seven features listed above are enregistered as modern “Yorkshire” dialect. This correlation also highlights the reliability of the use of textual data to investigate potentially enregistered features, as of the 11 features consistently listed as “Yorkshire” by the survey respondents listed in table 1 above, 7 of them were consistently discussed and frequently represented in the modern corpus data. Thus, it is likely that features meeting the criteria listed in (1)-(3) above in the nineteenth-century corpus data would have been similarly salient to nineteenth-century audiences and also potentially enregistered.

The study of enregisterment in historical contexts is therefore possible (see also Cooper 2014), as the patterns highlighted in the modern corpus data and the online survey suggest an enregistered “Yorkshire” repertoire. The correlations in the modern data indicate that textual data is indeed sufficient to discuss historical enregisterment, and we can consequently use the patterns in the data from the present to inform data from the past.

References


ADDY, S. O. (1888) *A Glossary of Words used in the Neighbourhood of Sheffield including a Selection of local names, and some Notices of Folk-lore, Games and Customs*, London: Published for the English Dialect Society by Trubner & Co. Ludgate Hill.


