Espaces Linguistiques

Investigating (in)coherence in Tyneside English(es): sociophonetic variation spectrums of FACE, GOAT, PRICE and MOUTH in the DECTE corpus

Analyse sociophonétique de la cohérence en anglais de Tyneside : les continuums variationnels de FACE, GOAT, PRICE et MOUTH dans le corpus DECTE

Maelle AMAND

University of Limoges, CeReS, UR 14922

maelle.amand@unilim.fr

Abstract: The present study investigates the co-occurrence of FACE, GOAT, PRICE and MOUTH vocalic variants in the Tyneside Linguistic Survey (TLS, Strang, 1968). The analysis is based on the original fine-grained transcriptions of the TLS sociolinguistic interviews across a sample of 44 predominantly working-class speakers (men/women, older, middle-aged and teenagers). Multiple Factorial Analysis and cluster analysis were combined to identify sub-groupings in the variety based on the entirety of the transcriptions (consonants, stressed and reduced vowels), and to investigate coherence across these vocalic variants. The results point towards major gender and class effects with four clusters of speakers characterised by varying degrees of accentedness. Overall, lower-status women and slightly higher-status women distinguish themselves through distinct variant choices. Differences amongst men of either status can best be measured through the extent to which they use the traditional variants.

Keywords: Tyneside English, coherence in sociolinguistics, aggregate approach, Multiple Factorial Analysis, DECTE corpus

Résumé : Cette étude propose une analyse de la co-occurrence des variantes phonétiques des ensembles lexicaux FACE, GOAT, PRICE et MOUTH (Wells, 1982) dans l'Enquête Linguistique de Tyneside (TLS, Strang, 1968). Les transcriptions phonétiques d'origine effectuées sur les dix premières minutes des entretiens sociolinguistiques servent de base à l'analyse, sur un échantillonnage de 44 locuteurs issus de milieux modestes (hommes/femmes, trois classes d'âge). Une analyse factorielle multiple (AFM, Escofier, 1985) suivie d'une classification a permis de classer les locuteurs en sous-groupes en fonction d'une utilisation plus ou moins marquée de l'anglais de Tyneside et d'analyser la cohérence des schémas variationels de ces groupes de locuteurs. L'intégralité du système phonologique disponible à travers les transcriptions a été pris en compte dans l'analyse. Les résultats indiquent que le milieu social ainsi que le genre sont des éléments déterminants dans la classification des locuteurs. Les femmes au statut social légèrement différent se distinguent par un choix plus systématique de variantes vocaliques. La distinction entre les hommes se mesure plutôt à travers la fréquence d'utilisation des variantes traditionnellement associées à l'accent de Tyneside.

Mots clés : Anglais de Tyneside, cohérence en sociolinguistique, approche par agrégat, Analyse Factorielle Multiple, corpus DECTE

Introduction

Tyneside English, henceforth TE, is a variety of English spoken in the north-east of England. The data reported here was collected from the Tyneside Linguistic Survey (Strang, 1968; Strang and Pellowe, 1969). It comprises a series of recorded one-on-one sociolinguistic interviews carried out in the two urban areas of Newcastle and Gateshead (Figure 1), the reading of a wordlist, and acceptability judgements (Allen *et al.*, 2007). Recordings were carried out in the late 1960s and early 1970s. One of the TLS's characteristics is that a portion of the interview material is matched with fine-grained phonetic transcriptions (Beal *et al.*, 2014). A sample of these transcriptions from the recorded sociolinguistic interviews of 44 out of 88 speakers from Gateshead and Newcastle is analysed in this paper–the remaining recordings and their matched phonetic transcriptions were not yet available at the time of the study.

Figure 1: map of Tyne and Wear, north-east England

Source: https://d-maps.com/carte.php?num_car=112548&lang=en

While most studies analysed one linguistic feature at a time (Trudgill, 1972), the TLS pioneered in opting for *aggregate* (Nerbonne & Wieling, 2017) approaches in language variation and change (Pellowe *et al.*, 1972), i.e. the investigation of the relationships among multiple linguistic variables. Nevertheless, the two approaches should not necessarily exclude each other, as Corrigan and colleagues highlight the complementary role of feature-based and aggregate approaches (Corrigan *et al.*, 2014) in revealing various aspects of co-variation in TE. This paper is a compromise between the two approaches: an aggregate statistical analysis is performed on highly enregistered (Agha,

2003) Tyneside vowel variants of FACE, GOAT, PRICE and MOUTH. Jones-Sargent (1983b) performed the first analysis on the fine-grained phonetic transcriptions of the TLS but was faced with the difficulty of a full aggregate analysis due to the computational limitations of statistical programs at the time. Moisl & Maguire (2008) analysed a less fine-grained level of transcription but flaws in the original hierarchical structure of the transcription coding itself led to a major bias amongst certain consonantal features, thereby potentially assigning greater weight to the vowels. The two studies, nonetheless, provided fundamental suggestions for future research: (1) reducing variational noise in the data, i.e. allotting less statistical importance to minor variation patterns, and (2) looking at how stable the clusters across various clustering approaches are. These two approaches were carried out by Moisl (2015) in a later study but with the same semi-fine-grained transcriptions. The present paper draws on the methodological approaches of the two studies: firstly, the use of the fine-grained transcriptions to limit biases caused by the hierarchical structure of consonantal features and secondly, that of noise reduction tools prior to classifying the speakers (cf. methodology section).

However, the innovative methodological aspect of this paper is the application of Multiple Factorial Analysis or MFA (Escofier & Pagès, 1984) to language variation, a tool generally used in sensorimetry (Husson *et al.*, 2001) or in the study of ecosystems (e.g. Lamentowicz *et al.*, 2010). MFA takes into account the hierarchical structure of variational data, namely, the link between a vowel and the variants pertaining to that vowel. It also reveals how linguistic variants covary. Namely, if a speaker uses a regional vowel variant for GOAT, do they use regional variants for the other three vowels? In addition, MFA summarises variation into trends through *dimensions* (ranging from broad to subtle)[1], without obliterating individual variation.

MFA cannot draw boundaries between individuals to form sociolinguistic groups, which is often relevant in language variation. However, when coupled with clustering analyses (Husson *et al.*, 2010), groups with similar variation patterns can emerge. The combined approaches are therefore particularly suited for sociolinguistic research.

Following British sociophonetic conventions for the description of English varieties, the standard keywords FACE, GOAT, PRICE and MOUTH, proposed by Wells (1982, p. xviii-xix), are used to represent the lexical sets containing the four closing diphthongs with the following British Received Pronunciation citation form /eɪ/, /əʊ/, /aɪ/ and /aʊ/. In Tyneside English (TE), words from the

---

[1] Dimensions are also known as factors in a factorial analysis, or principal components in a principal component analysis.

PRICE set often alternate between a raised onset [ɛɪ] and a low onset [aɪ]. This alternation is partially conditioned by what follows the vowel (Milroy, 1996). Milroy's analysis revealed that working-class speakers favoured a raised onset [ɛɪ] in PRICE and that younger male speakers appeared to lead the swing towards a less socially restricted use of this raised variant (Milroy, 1996). Monophthongal realisations for MOUTH [uː] are highly enregistered in Tyneside (Beal, 2004)–i.e. the general public is aware that this vowel is typical of Tyneside. But for most speakers, it is often used for a limited number of words like *town* or *brown*, respectively spelt *toon* (the *Toon Army*)[2] and *broon* (*Westmorland Broon ale*), to reflect the local pronunciation. Beal (2004) states that the most widespread realisations of MOUTH in Tyneside are closer to [ɛʊ].

FACE and GOAT are, to some degree, partner vowels working in *lockstep* since "the ways in which they vary in terms of their phonetic exponency and the distribution of these variants across the TE-speaking population seem to work in parallel" (Watt & Foulkes, 2017, p. 157). The most commonly used variants are [eː] and [oː] for FACE and GOAT, respectively (Watt, 1998). They are said to be unmarked in the area because they "accommodate a less marked identity as Northerners" (Haddican *et al.*, 2013). They will be referred to as the *pan-northern variants* (*Ibid.*). The local variants are often realised as forms approximating [ɪə] for FACE and [ʊə] or [øː] for GOAT (Corrigan, 2012).

Closing diphthongs [eɪ] and [oʊ] are often heard among middle-class speakers (Watt & Foulkes, 2017) and will be call *prestige-variants*. Given the asymmetry across the variant inventories for each vowel and the social indexicality observed in more recent TE speech (Watt, 1999; Milroy, 1996), the working hypothesis of this paper is that coherence in TE is only partially the result of linguistic constraints such as the Scottish Vowel Length Rule (Milroy, 1996) or the type of phonemes on either side of FACE and GOAT vowels (Buchstaller *et al.*, 2017) and that social factors, first and foremost, play a major role in accounting for correlation in usage–as found in many sociolinguistic studies on other varieties (starting from Labov, 1972; Macaulay 1977 and Trudgill 1974). However, studies investigating co-variance (*coherence*) in a variety are rarer (cf. Guy, 2013; Beaman & Guy, 2022). This was also amongst the original aims of the Tyneside Linguistic Survey (Jones-Sargent, 1983, p. 21-22). Following De Camp's study in 1971 on Jamaican English, more recent approaches have delved into the blurry boundaries of sociolinguistic stratification, thereby looking at

---

[2] The local football club.

continuums instead of discrete groups (e.g. Amand, 2019 on TE, Sewell, 2022 on Hong Kong English consonants).
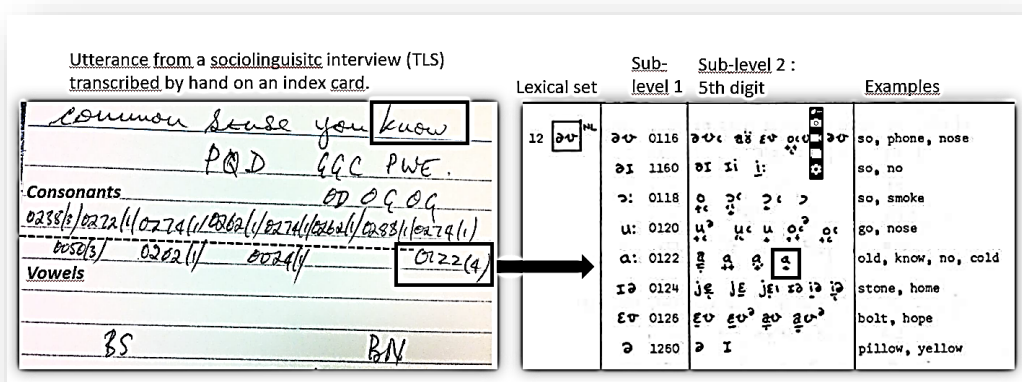
The purpose of this paper is to report intercorrelations of four highly enregistered (Agha, 2003) TE vowel variants, i.e. those of FACE, GOAT, PRICE and MOUTH based on the original TLS phonetic transcriptions. The following questions are raised: (1) **Lects**: how many linguistic groups can be found through the variant choice of FACE, GOAT, PRICE and MOUTH only? (2) **Coherence**: do these groups use variants consistently across FACE, GOAT, PRICE and MOUTH in TE? (3) To what extent are the lects, and the coherence within, socially motivated? (4) Can variation patterns be defined less in terms of discrete lects, but rather in terms of multiple variational **spectrums**?

1. Methodology: the corpus

1.1. The TLS phonetic transcription scheme

The data analysed in this paper is referred to as the TLS-coding data. It is composed of fine-grained phonetic transcriptions (carried out by trained phoneticians) of the first ten minutes of recorded sociolinguistic interviews. The recordings were made in the participants' homes. These transcriptions comprise consonants, stressed and reduced vowels. To make the transcriptions computer readable, each phonetic variant deemed likely to appear in speech was encoded as a sequence of five digits (Jones-Sargent, 1983, p. 295-302), incorporating a two-level hierarchical structure (Figure 2). The full coding-scheme is available on the DECTE website (Jones-Sargent, 1983a)[3].

Figure 2: original transcription cards (left) & TLS coding-scheme (right)

---

[3] https://research.ncl.ac.uk/decte/transscheme.htm

1.2. Phonetic data

The sample of the TLS-coding data analysed is composed of frequencies of 566 variant types distributed across all 63 phonetic features (consonants and vowels) during the first 10 minutes of a sociolinguistic interview for 44 speakers. The number of variant types across the 4 vowels amounts to 84. The total number of transcribed variants amounts to 105,204 tokens, 9,428 of which pertain to the FACE, GOAT, PRICE and MOUTH lexical sets. The median number of occurrences per speaker for the latter four features ranges between 21 and 57. Frequencies were computed from raw counts of all variants within each lexical set.

Table 1: counts per speaker per vowel (total count: top row, speakers: n= 44).

SUM FACE	1 705	SUM GOAT	2 515	SUM PRICE	4 186	SUM MOUTH	1 022
SPEAKER MEAN	39	SPEAKER MEAN	57	SPEAKER MEAN	95	SPEAKER MEAN	23
SPEAKER MEDIAN	36	SPEAKER MEDIAN	57	SPEAKER MEDIAN	99	SPEAKER MEDIAN	21
MIN	16	MIN	11	MIN	19	MIN	6
MAX	75	MAX	107	MAX	151	MAX	55
Q1	29	Q1	46	Q1	82	Q1	17
Q3	48	Q3	70	Q3	110.75	Q3	29

1.3. Selected social data

The original aim of the TLS was to include as many social variables as possible–father's occupation, parents' birthplace, political affiliation, leisure, identification with the area, and so on, in order to build a hierarchy of the social characteristics that were more significantly linked to the linguistic variational patterns (Jones-Sargent, 1983). In the present study, we chose to retain macrosociological categories (Eckert, 2012) such as age, gender and social class. A fourth category regarding whether speakers had continued their education after the minimum school leaving age was also included.

1.4. Participants

We selected the 7 Newcastle and 37 Gateshead participants from the original TLS study (see Table 1) whose data had been checked by the DECTE team, and made available on the DECTE website as of 2014. Since these speakers were randomly selected from the Electoral Register of 1966 by the "rateable value per dwelling by polling district" (Pellowe *et al.*, 1972, p. 23-24), the number of participants within each macrosociological category, i.e. age, gender and class, is not necessarily even across all categories. The class categories are based upon the speaker's (former) occupation. The *Classification of Occupations* of 1966 served as a reference to better fit with the social categories defined by the original TLS team in the 1960s-70s.

Table 2: Social background of the 44 TLS informants

Lower-middle class[1]	Women (n=6)	Men (n=5)
	Civil servant, school secretary, nurse, employment exchange worker...	Skilled workers (engraver, cast operator), teacher trainee...
Young (F=3, M=3)	G05F, G09F, G21F	G20M, G23M, N04M, N02M, N06M
Middle-aged (F=2, M=2)	G11F, G34F, N01F, N03F	--
Older (F=1, M=1)	--	G28M

Working class[2]	Women (n=15)	Men (n=13)
	Home help, factory worker, cleaner, waitress, shop assistant...	Painter, plumber, storeman, miner...
Young (F = 3, M = 2)	G08F, G10F, G15F	G12M, G27M
Middle-aged (F=12, M=10)	G01F, G03F, G16F, G17F, G22F, G25F, G26F, G32F, G35F, G36F, G37F, N05F	G02M, G04M, G07M, G14M, G18M, G19M, G29M, G31M, N06M, N07M(?)[3]
Older (F=1, M=4)	G06F	G04M, G13M, G24M, G33M

Middle-aged speakers predominate. We therefore expect a certain degree of *retrenchment* (Chambers, 2003, p. 195; Wagner, 2012, p. 375) from the use of more local or less-standard features used in youth. Two speakers were not born in the North-East (G31M: Northern Ireland, N06M: south-east London) but were included in previous studies (Moisl, 2015; Jones-Sargent, 1983b). In our study, these *outliers* serve as points of reference in the statistical analysis to highlight the distance between native Tynesiders and speakers who have settled in the area but were not born and raised in the North-East. They contribute to building a limit in the variational space computed by the statistical analysis, whose first aim is to simplify complex data and show macro-groupings of individuals (native Tynesiders vs. non-native Tynesiders) before going into meso-groupings within the native Tynesiders, e.g. broad vs. less broad.

The next section details the statistical methods applied to the corpus.

2. Methodology: statistical analysis

2.1. Multiple Factorial Analysis + clustering in language variation

The statistical tool used here to analyse the data is rarely used in sociophonetics. The methodology of the approach along with the interpretation of the results will therefore be detailed in greater length than is commonly found in articles on language variation.

A Multiple Factorial Analysis was carried out with the R package {FactoMineR} (Lê *et al.*, 2008). It was selected over other statistical methods because it is *feature-centric* (Amand, 2019). Namely, each variant (or *exponent*) is not treated statistically as an independent variable but as a variable that is hierarchically linked to its pertaining phoneme or *feature*[4]. Hence, the results are first apprehended from the perspective of the *feature* itself, e.g., how variation in FACE helps define speakers socially.

Unlike more common methods such as Principal Component Analysis (PCA), MFA considers the structural relationship between the linguistic features and their variants. Since there are over 20 different variants reported for FACE and GOAT (see Amand, 2019), a *variant-centric* approach with PCA would have generated output that is difficult to interpret because each variant is treated as an independent variable and not as pertaining to a lexical set. The approach is similar to the one used in Ghyselen and De Vogelare (2018) who computed speaker-profiles based on correspondence analysis, which also allows for sublevels of a main variable.

Another reason for using MFA is that it gives less frequent vowels like SQUARE (RP /eə/) or vowels with more complex variation "a chance to have the same importance in the final analysis" (Picardi *et al.*, 2020) as it balances out the role of each lexical set in accounting for variation. In a canonical PCA analysis, features with binary variation such as *-ing* (velar /ŋ/ or apical /n/) will appear as overly important in the sociolinguistic discrimination of speakers. Namely, the easier it is to split speakers (those who favour /ŋ/ vs. /n/), the more salient the lexical set will appear in the analysis. The other features, with more complex variation, will be minimised (Escofier & Pagès, 1994). Since the number of reported variants are not the same across the four vowels under scrutiny, MFA was deemed a more suitable approach than PCA.

A cluster analysis was then performed based on the MFA output. This provides an ideal combination: (1) MFA prunes variational noise (i.e. a variant which occurs only once across two

---

[4] The term feature is used here because the approach is not restricted to phonetic variation and can be applied to other types of variation (lexical, grammatical etc.).

speakers) without downplaying the role of relevant variation patterns, and (2) the cluster analysis builds on the MFA analysis to identify more stable sub-groups of speakers (Husson *et al.*, 2010). The Ward method, often also called the 'minimum variance method,' was used for clustering since it has proven relevant for the TLS data analysis (Moisl, 2015) and in several other linguistic studies (Deumert, 1999; Gries, 2009). It also minimises intra-cluster variance (Janssens *et al.*, 2008; Gries, 2009) so as to limit the number of groups containing one or two speakers (see Amand, 2019 for a comparison of clustering methods on the full TLS data). Finally, I evaluated whether these groups shared social characteristics using chi-square tests provided by {FactoMineR}.

2.2. Preliminary analyses: assessing the robustness of the MFA results

Preliminary analyses confirmed the lack of randomness in the data (Husson *et al.* 2011, p. 229) in variation and the need to favour MFA over other methods like PCA (see Amand, 2019, p. 345 ff.). The MFA analysis was first performed on the dataset with all 566 phonetic variants from the 63 phonological features (consonants, full vowels and reduced vowels), so as to check whether the variants of FACE, GOAT, PRICE and MOUTH have comparable discriminating roles within the context of the original 63 phonological features retained in the TLS coding-scheme (see Amand, 2019, p. 355).

In MFA, factors, aka. *dimensions*[5], have to be viewed in terms of axes with opposite variational patterns at each end. Another element to consider is the presence of diverging variant usage, with specific variants emerging as representative of opposite variant usage. These variants play a major role because they define the limits of the variation spectrum which corresponds to a dimension. The first five variables that contribute the most to the definition of the first two dimensions percentagewise are examined. The higher the score, the better the variants help define opposite trends between speakers. Given the high number of variables included in the model, these percentages are necessarily low. But if the 84 variants had contributed to the dimensions evenly, the percentage would have been 1.19%, i.e. 84/100, while the top contributions are between three to eight times as high. Contributions were deemed reliable enough.

Another rule of thumb is to check whether eigenvalues are superior to 1. Here, eigenvalues drop below 1 from the sixth dimension onwards, (cf. Baayen, 2008) and the first six dimensions have eigenvalues ideally superior to 1 (Di Franco & Marradi, 2013), which means that the first five dimensions are relevant in measuring variation across the sample of speakers. The remaining ones

---

[5] In the "French School of PCA" (Holmes, 2006), factors are more generally called dimensions.

are potentially noise or marginal variation patterns. However, only the two dimensions, with respective eigenvalues of 2.81 and 1.86, will be used to plot the clusters on the factor map, since visualising six dimensions simultaneously leads to a reduced comprehensibility of the results (Ghyselen & De Vogelaer, 2018). After an inspection of the clustering output results, a maximum of nine groups was retained for the classification process because a lower number of groupings would have ironed out sub-groups using rarer variants.

3. Results

3.1. An overview of the main variation patterns

Aggregate percentages (Figure 3) provide a general overview of dominant forms in Tyneside in the 1970s in the given sample of 44 speakers. In FACE and GOAT, the pan-northern monophthongal realisations in yellow predominate (FACE: 59%+2%=61%, GOAT: 44%+7%+3%=54%). Prestige forms with closing diphthongs (in blue) have similar scores (respectively 10%+3%+2%=15% vs. 16%+2%=18%), so do localised variants (17% vs. 21%). GOAT and MOUTH have a wider range of localised variants (green) with various monophthongal and diphthongal variants. The highly enregistered monophthong for MOUTH (as in *Toon*) is indeed extremely rare in casual interview speech (7%).

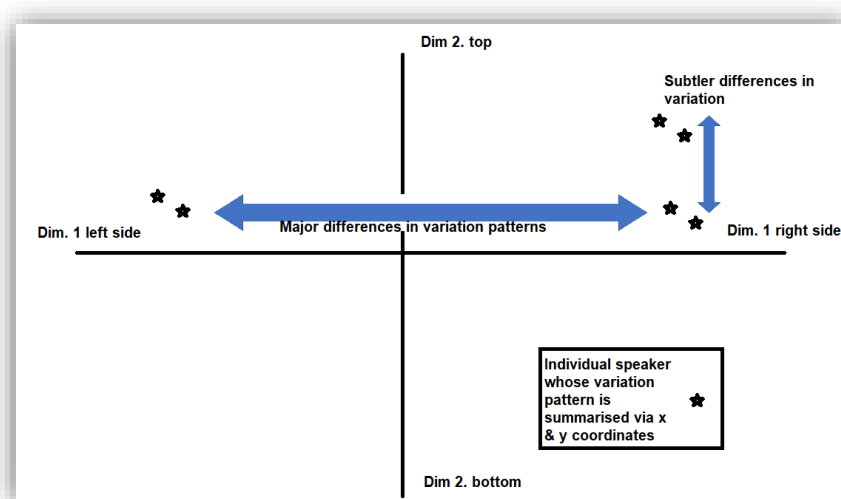Figure 3: percentage frequency per variant per vowel for all 44 speakers

28 speakers use the pan-northern variant [eː] for FACE over half the time, but only 20 use [oː] for GOAT above 54% of the time, this means other variants compete with the pan-northern one. 17 speakers use the raised onset [ɛʊ] in MOUTH above 50% of the time. Scores for the raised onset [ɛɪ] in PRICE are lower due to the inclusion of the personal pronoun *I*, often realised as [ɑː] (Moisl & Maguire, 2008) but 20 speakers use the raised onset around three times as much as the low onset.

3.2. "Are yi broad?": speakers at opposite ends of 2 variation spectrums

Interpreting the dimensions: here, dimensions reflect spectrums where speakers are placed depending on their overall use of variants across the 4 vowels. Interpretation can be made easier through a plot with two axes crossing at a 90-degree angle (Figure 4). Speakers with completely different variation patterns are placed at opposite ends of that spectrum. Dimension 1 generally displays speakers on an obvious opposition: those who use non-local vs. those who use local forms. Dimension 2 shows a slightly subtler pattern opposing speakers at either end. Namely, amongst those who do not use standard forms, sub-variational trends emerge: users of vestige forms vs. users of pan-northern forms. Other dimensions show micro oppositions amongst the pan-northern speakers or amongst the traditional speakers.

Figure 4: Schematization of the first two dimensions of MFA, used to highlight major and subtler variation patterns in sociophonetics

Oppositions in variant use through dim. 1 & 2: the MFA diagnostics for dimension 1 (Table 2) highlight an opposition between the speakers who exhibit a very homogeneous variation pattern in all four vowels (the more socially prestigious forms), and the speakers who use the regional variant [ɛɪ] in PRICE, but who do not necessarily use the same variants in FACE, GOAT and MOUTH.

This means that there are two main groups: one who uses standard forms consistently across the 4 vowels, and another group whose speakers commonly favour [ɛɪ] in PRICE significantly, but who distinguish themselves through the use of distinct regional forms. The opposition within the latter group is reflected in dimension 2, which reveals an opposition between the speakers who use the supralocal variants [eː] and [oː] in FACE and GOAT, plus the raised onset in MOUTH ([ɛʊ]), and those who exhibit the traditional variants of PRICE [i] and MOUTH [u].

Table 2: Top 5 contributions of the phonetic variants to dimensions 1 (top) and 2 (bottom) along with the position of the variants on the dimension axes when variant use is above average

Dim. 1		Above supralocal vs. regional speakers	
Variants at opposite ends of dimension 1 (top contributions in %)			Side of the axis on factor map when score is high
PRICE	[ɛɪ]	7,83	Right (positive coordinates)
FACE	[eɪ]	5,19	Left (negative coordinates)
GOAT	[oʊ]	4,97	Left
PRICE	[aɪ]	4,6	Left
MOUTH	[aʊ]	4,08	Left

Dim. 2		Supralocal vs. traditional speakers	
Variants at opposite ends of dimension 2 (top contributions in %)			Side of the axis on factor map when score is high
GOAT	[oː]	9,77	Bottom (negative coordinates)
FACE	[eː]	6,22	Bottom
MOUTH	[ɛʊ]	5,97	Bottom
MOUTH	[u]	3,71	Top (positive coordinates)
PRICE	[i]	3,3	Top

Summary of both tables

Dim. 1: (a) if you had to find one vowel that distinguishes typical Tyneside speakers from pan-northern speakers, take PRICE and look for the raised onset; (b) when a speaker uses pan-northern forms, they seem to use them across the four vowels; (c) within the typical Tyneside speakers, there is a high degree of variation and sub-groups can be found through dimension 2.

Dim. 2: (a) across typical Tyneside speakers two trends stand out: speakers with high scores of pan-northern forms in GOAT, FACE and MOUTH vs. speakers with vestigial forms in PRICE and MOUTH.
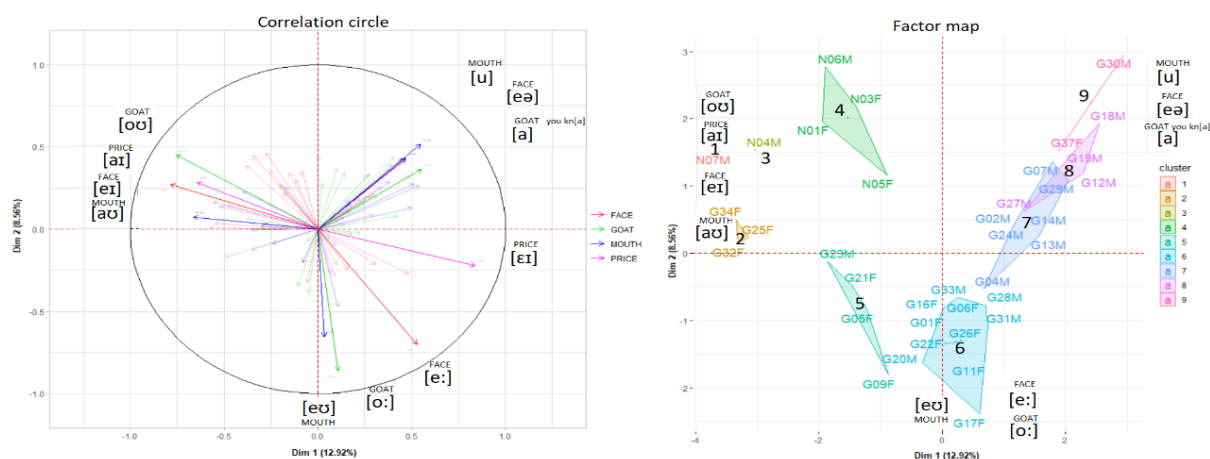
However, these covariational trends need to be confirmed through an inspection of the correlation between the variants, and groups of speakers, confirmed via a cluster analysis. This is dealt with in the next section.

3.3. Coherence across the 4 vowels and across speaker groups

Covariation patterns will now be examined in more depth via a correlation circle displayed in Figure 5 (left). Correlation circles are particularly useful for visualizing covariation amongst a high number of phonetic variants in two dimensions simultaneously. Variants are depicted in the form of arrows. The closer the arrows are to the rim of the circle, the better they reflect a clear opposition between speakers in one or both dimensions (Abdi and Williams, 2010). To ensure a better interpretability of the graph only the 12 phonetic variants with the highest contributions to dimensions 1 and 2 are labelled here[6]. Arrows forming acute angles are positively correlated (FACE [eə] & MOUTH [u:]). Obtuse angles indicate a negative correlation (raised vs. low onsets in PRICE). Variants with no correlation form a right angle (GOAT [a] vs. FACE [e:]).

Three main covariation patterns emerge: the top left part of the circle displays covariation amongst high prestige variants and the top right, amongst traditional ones. The bottom part of the graph shows a more supralocal type of variation with a raised onset in MOUTH and the pan-northern [e:] and [o:] in FACE and GOAT. The arrow for the raised onset in PRICE [ɛɪ] is placed in between the pan-northern and traditional arrows. This means that many speakers from both groups use this raised onset.

Figure 5: correlation circle (left): coherence in variants & factor map (right): individual speakers their affiliation to a cluster

The correlation circle confirms the output for dim. 1. Prestige variants covary: speakers who use closing diphthongs in GOAT and FACE, use low onsets for both PRICE and MOUTH. The traditional

---

[6] There were two monophthongal variants for MOUTH but [u] was used to label them both (TLS code: 01462 and 01464).

variants that covary in a highly systematic way are the centring diphthong [eə] in FACE and two monophthongal realisations in MOUTH ([u]). The variant [a] for GOAT is also positively correlated with the traditional variants [eə] and [u]. The raised onset in MOUTH [ɛʊ] is strongly correlated with the pan-northern variant [oː] for GOAT. The pan-northern FACE variant [eː] is less strongly correlated with [ɛʊ] and [oː] because many traditional speakers use it too. This is also the case for the raised onset in PRICE. The correlation circle, however, does not show individual speakers and limits itself to coherence amongst variants. Another graph with the same coordinates needs to be added to reveal the distances between each speaker.

The right part of Figure 5 illustrates the distances between individuals and is called a *factor map*. This graph is the counterpart of the correlation circle and the two should ideally be interpreted side by side since they represent the same spectrums represented by dim. 1 and 2, albeit highlighting different aspects of the data, which is similar to the use of different filters on the same map. While the correlation circle highlights coherence within the variants chosen by speakers, the factor map places the speakers based on their overall variant usage. On the factor map, speakers with high percentage scores in one of the above-mentioned variants will be placed around where variants' arrows are pointing in the correlation circle. Speakers with similar phonetic variation have similar coordinates (G22F and G26F). By contrast, G32F, G17F and G30M have very distinct pronunciation patterns. The closer the speakers are to where axes cross zero, the closer they are to the sample's average variation pattern, e.g. G33M and G04M.

Polygons circumscribe speakers that were grouped together in the cluster analysis. The coordinates of certain speakers suggest considering at least more than 6 groupings and setting apart atypical speakers that are placed far away from the rest of the sample (N07M or N04M). We now comment on the results of the cluster analysis.
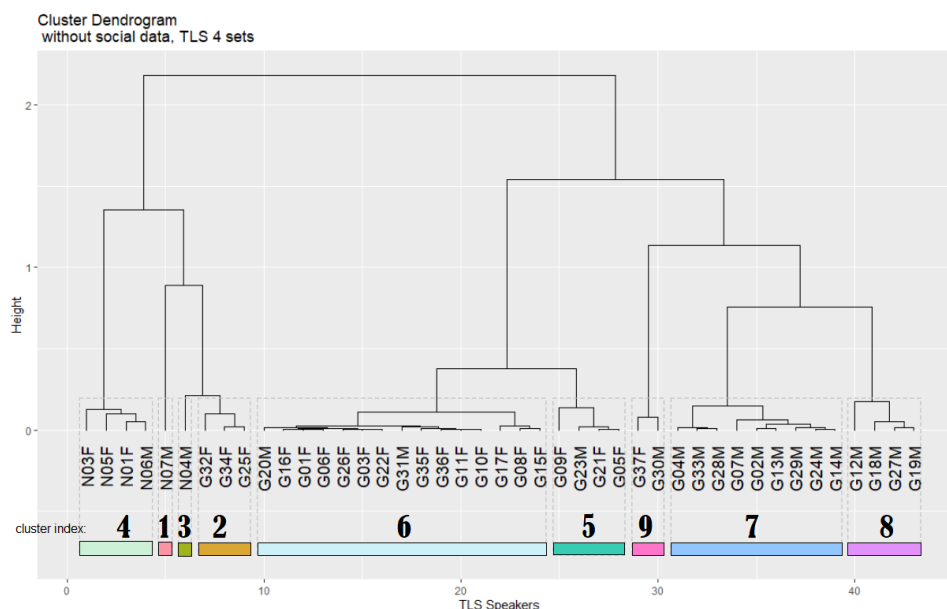
3.4. Sub-linguistic norms in Tyneside English?

How many dimensions retained for the clustering analysis? In order to analyse the clustering tendencies in more depth, the first five dimensions of the MFA were used as input for a cluster analysis–the eigenvalue for the sixth dimension was 1.01 and was disregarded as it was too close to the rule of thumb threshold of 1. The analysis was consolidated with k-means, so as to optimise group homogeneity (Lebart *et al.*, 2000). After inspecting the sum of squares, it was decided to stop at the last large drop in units, namely, between 8 and 9 clusters, in order to have a better idea of

the variational cohesion within and between these sub-groups. Going beyond that point would have led to overfitting (cf. Baayen, 2008).

Figure 6: classification of speakers. First letter of speaker index: N = Newcastle, G = Gateshead. Last letter: F = woman, M = man. Groupings are the same as Figure 3. Y-axis: the higher the node, the more different the overall variant patterns. X-axis: cumulative sum of speakers (n=44)

What are the linguistic groups of speakers? Figure 6 illustrates the dendrogram resulting from the classification. It indicates that the speakers are divided into two main categories (cut-off point superior to 2, y-axis): the Newcastle speakers and all but three of the Gateshead speakers (G32F, G34F and G25F). The higher the split, the larger the difference between the speakers and the lower the split, the closer the speaker's values approach the cluster's average scores. Based on the MFA plots (Figure 3), we know that this reflects the opposition between the speakers with a higher share of prestigious variants and those exhibiting higher rates for the regional variants (dim. 1). The second major split (cut-off point of 1.5) is the one between the more traditional speakers and those who predominantly use pan-northern variants (dim. 2). Beyond oppositions defined in dim. 1 and 2, the speakers who use the prestigious variants can also be sub-divided into two categories, with N07M being considered an outlier. **There are also three categories of traditional speakers and two for pan-northern ones**. We now examine which clusters are correlated with one or more social factors.

How do these groups correlate with social factors? Chi-square tests provided in the MFA + clustering output tables assess the link between the linguistic groupings and the social factors, here: class, age, gender, education and sampling area (Newcastle vs. Gateshead cohorts).

The selected percentages given in this section are also found in the raw output tables. With the default level of group cutting, (i.e. 6 sub-groupings) one major effect is the city of residence (p = .009). This is partly due to variations in phonetic transcriptions[7], and, potentially, to class[8]. Yet, the social factor that best helps discriminate speakers is gender (p = .003). This effect is clearly visible in Figure 4, with most speakers being grouped by gender. This suggests that in Tyneside English, there are distinct male and female linguistic norms reflected in the phonetic variants of FACE, GOAT, PRICE, and MOUTH. With the 9 sub-groupings[9]: class, sex and city of residence significantly contribute to define at least one of these 9 groups (respectively p. < 000003, p < .003 and p < .03). Let us now examine what characterises these groups from a socio-linguistic point of view.

Cluster 6 includes more speakers (n=14). It is predominantly female (p < 0.05, 13 women vs. 2 men) who did not pursue further education beyond minimum school leaving age (p < 0.05, 88.2% with no further education). This group extensively opt for the pan-northern variants [oː] and [eː] for GOAT and FACE and favour the raised onset in MOUTH [ɛʊ] (above 75% for the three variants). Both local *and* high prestige variants are avoided in this group as they remain below 6%.

Pan-northern female speakers → monophthongs for GOAT and FACE, raised onsets for PRICE and MOUTH. → highly coherent across the 4 vowels.

By comparison, cluster 5 is exclusively composed of younger speakers from lower-middle class background and with further education (class: p = 0.003; education: p = 0.004; age: p = 0.004)[10]. Clusters 5 and 6 mainly differ in variant usage for MOUTH and PRICE. While women with no further education in cluster 6 favour the raised onsets for both these vowels, the group of younger lower-middle class speakers used low onsets above 90% for MOUTH (except for G09F who uses the raised onset more than 60%). However, both groups predominantly use the pan-northern variants [oː] and [eː] for GOAT and FACE (except for G21F who has an equal share of prestige and pan-northern variants in FACE).

---

7 There were two distinct phonetically-trained transcribers from Newcastle University. Although both followed the same transcription schemes, minor differences in fine-grained transcriptions were observed. A more careful examination of these differences will be carried out in a future study.

8 This issue is being investigated further since more information is being made available on the Newcastle speakers.

9 The clusters are consolidated with k-means. This parameter has an impact on how clusters are made and on the final diagnostics.

10 Although the group includes a majority of women (3 women vs. 1 man), the group is too small for gender to be deemed a significant effect.

Lower-middle class younger speakers → monophthongs for GOAT and FACE, raised onsets for price but low onset for MOUTH. → non-localised speech in MOUTH only.

Clusters 7, 8 and 9 are almost exclusively composed of men, except for G37F in k9. Clusters 7 and 8 pattern in a similar way for FACE (37.1% vs. 38.3%) and PRICE (37.6% VS. 44.0%) but cluster 8 is characterized by the significant, albeit infrequent use of the vestigial variant [u] in MOUTH (18.3%).

In addition, two separate retracted MOUTH variants seem to define these two groups–cluster 7 [ɔʊ] 45.0%, cluster 8 [ʌʊ] 34.5%. An acoustic analysis along with a perception test should be carried out to assess the differences between the two variants. Cluster 9 groups two atypical speakers (G30M, G37F) who, combined, use [u] for MOUTH twice as much as cluster 8 (53.0% vs. 18.3%). These speakers also used significant amounts (ca. 11%) of rare variants like [ɪi] and [ɛʊ] for GOAT, and [ɪʊ] for MOUTH.

Speakers using highly localised forms are predominantly men (except for G37F). But some speakers distinguish themselves through the use of [u] in MOUTH (к8 & к9), while others use retracted onsets [ɔʊ] or [ʌʊ] (k7 & k8). Working-class men form a continuum from broad to very broad (k7 → k8 → k9).

Cluster 4 is exclusively composed of speakers from the Newcastle sample. It is defined by particularly high scores of [oʊ] in GOAT (65.6%) and two competing variants for FACE (pan-northern 11.5% vs. prestige form 17.8%). It has a large share of raised onsets in MOUTH (43.3%) and 10% of low onsets for PRICE. **Coherence in this group is less obvious because it is in between two linguistic profiles: pan-northern speakers and those using prestige forms.**

By opposition, cluster 2 uses the low onsets in MOUTH and PRICE above 80%. This group also used the prestige variant [eɪ] in FACE more extensively than cluster 4 (k2: 74% vs. k4: 17.8%). It does not quite work in lockstep with GOAT ([oʊ] in 25%). Cluster 2 is only composed of three Gateshead women and none of the selected social variables significantly characterize the cluster due to its size. We hypothesise that **cluster 2 forms the middle part of the female variation continuum ranging from non-localised (k4,1,3) to female pan-northern (k6 & k5).**

4. Discussion

In the Gateshead sample, variation is highly determined by gender: men use variants co-occurring amongst men only. The same is found for women, with a few rare exceptions (G37F, G20M). The

broad⇔less broad spectrum for Gateshead men ranges from pan-northern forms to vestigial forms. Women's spectrum (except for G37F) goes along the line of pan-northern⇔prestige forms. As expected, the common ground is pan-northern features used by both men and women, which is supported by the work of Haddican et *al.* (2013).

In prestige and pan-northern variants for FACE and GOAT, co-variation is coherent whereas, for local variants, co-variation was harder to assess due to the wide array of variants reported for GOAT compared to those for FACE, which was mainly [eə]. In addition, the more frequent use of pan-northern [e:] for FACE amongst the traditional speakers indicate a looser symmetry between FACE and GOAT in the masculine and traditional sub-varieties of TE.

PRICE and MOUTH also co-varied tightly amongst most speakers using supralocal and high-prestige variants for FACE and GOAT. While users of supralocal forms for FACE and GOAT favour a raised onset for both PRICE and MOUTH, speakers using high-prestige variants for FACE and GOAT can be divided into two sub-groups, i.e. those who use low onsets for both PRICE and MOUTH and those using a low onset for MOUTH but a raised onset for PRICE. The latter group corresponds to younger lower-middle class women who benefited from further education. Amand (2019, p. 513 ff.) investigated the interactional effect of class and age for the raised onset in MOUTH, which confirmed that, in the 1970s, the variant simultaneously indexes class and age in TE–young working-class women use the variant twice as much as their lower-middle class female counterparts (respectively > 80% vs. > 40%). However, this variant did not seem to index class amongst older women whose scores range between 70% and 80%. It is beyond the scope of this paper to elaborate upon age effects but it is possible that this age effect implies standard change with the low onset in MOUTH becoming the new local standard amongst women of both classes as the latter variant is more often used in read speech in the TLS and even more so in the read material of a 1990s corpus of TE (Amand, 2019). It is also possible that the low onset in MOUTH is simply an extension of the pan-northern coherence observed in GOAT and FACE to the mouth vowel.

Raised onsets in PRICE co-vary in a peculiar way since its covariation with pan-northern variants in GOAT and FACE and the raised onset in MOUTH does not prevent a simultaneous co-variation with the traditional variants of FACE, GOAT and MOUTH. This indicates that in the 1970s, a raised onset in PRICE indexed class more than it did gender. These results are in line with Milroy's study on a later corpus of TE (Milroy, 1996). It may also index an intermediate degree of localness encompassing both users of traditional forms of TE in other vowels and speakers with a more levelled form of TE in their use of pan-northern variants in FACE and GOAT.

One regular pattern observed here is that the working-class women of the sample systematically score higher in the pan-northern variants of FACE and GOAT, which is in line with Watt's study of conversational material from the TLS corpus (Watt, 1998). They nonetheless retain raised onsets in PRICE and MOUTH; the former indexing class or localness, while the other seems to index both class/localness *and* gender. In addition, women form more coherent sociolects than men because their choice of variants for each vowel is more limited, which converges with Guy's findings that women, contrarily to men, "consistently show co-varying usage of socially stratified variables" (Guy, 2013, p. 69). Nonetheless, men are also coherent in the sense that they systematically avoid the variants adopted by the less traditional women.

Interestingly, G20M, is an atypical speaker concerning MOUTH. He was the only man using the record score of 90% for the raised onset in MOUTH vs. an average of 36% amongst the other men. Buchstaller *et al.* (2017) analysed his realisations for FACE 42 years later, in a second sociolinguistic interview, in similar conditions to the original TLS interview. At the time of the first interview, he was a teacher trainee in his twenties and his FACE variants were almost categorically monophthongal. Given the increasingly high number of women trained as teachers in the 1970s (Worth, 2019)[11], It is possible that G20M used the raised onset in MOUTH as a form of accommodation to his colleagues at work (cf. Coupland 1980 on style shifting in the work-place), a habit that was triggered by the semi-formal setting of the sociolinguistic interview, also known as the *observer effect* (Labov, 1972 and 1981). In the second interview, he had experienced considerable upward mobility and his realisations of FACE shifted towards "the most standard repertoire in the panel sample" (Buchstaller *et al.* 2017, p. 24), namely, the closing diphthong [eɪ]. G20M works in a language sensitive occupation (Chambers, 1994), in which the "legitimised language" plays an important role in shaping variation patterns amongst speakers, which may account for this remarkable shift. It is hypothesised that this change in FACE in favour of a more southern-like form is likely to be observable in MOUTH and PRICE too, with a complete retrenchment from the use of the raised onsets for both vowels, and higher ratios of closing diphthongs for GOAT. A future study on coherence in FACE, GOAT, PRICE and MOUTH on the second interviews of the six original 6 TLS speakers may provide insight as to how coherence evolves across the lifespan.

---

[11] The 1970s were a turning point for female educators. As Worth explains, "there was a marriage bar in place across much of the public sector until the 1940s and, crucially, part-time or more flexible roles were not widely available in the welfare professions until the 1970s". She adds that "in the 1971 census, 381,000 women were studying under subjects classified as 'education', compared to only 140,000 men".

Conclusion

This paper explored sociolectal cohesion within Tyneside English. In the Gateshead sample, four groups emerged as being sharply stratified by gender and social status. Although, at first sight, women seem more coherent in their variant use, men are also coherent in their own way. Women may treat socially stratified variants as discrete entities, i.e. opting for either one variant or another distinct one, whereas men seem to treat socially stratified variants as a continuum, thereby distinguishing themselves through the degree in which they exhibit traditional variants.

Although certain linguistic groups seemed less coherent than others with regards to the four vowels under scrutiny, the average standard deviation for groups of more than two speakers ranged between 6 and 11%. The holistic model, "in which particular varieties or registers characterised by clusters of socially evaluated variables become identifiable in a community" (Guy, 2013), is supported by the fact that a certain number of speakers have MFA coordinates that approximate the centroids of a cluster in a factor analysis. They are what Pagès (2013) calls *paragons*, or the speakers with the average characteristics of the cluster they belong to, e.g. the cluster of "Geordie[12]" men: G14M, G13M, and G24M (respectively, former factory worker, former miner and former wire drawer, aged between 40 and 60). This means that these paragons' variation patterns can form the basis of holistic representations of sociolinguistic sub-groups, with, at the periphery of clusters, innovators of sound change, social risers (G20M, the teacher trainee) and speakers that keep vestigial forms alive (G19M, the JCB driver or G37F, the factory worker).

References

ABDI Hervé & WILLIAMS, Lynne. J. 2010. "Principal component analysis", *Wiley interdisciplinary reviews, Computational statistics*, vol 2, p. 433-459.

AGHA Asif, 2003, "The social life of cultural value", *Language and Communication*, vol 23, p. 231-73.

ALLEN Will, BEAL Joan C., CORRIGAN Karen. P., MAGUIRE Warren. and MOISL Herman L., 2007, "A Linguistic 'Time Capsule': The Newcastle Electronic Corpus of Tyneside English", in BEAL Joan C., CORRIGAN Karen P. & MOISL Herman L. (dir.), *Creating and digitizing language corpora: Volume 2: Diachronic Databases*, London, Palgrave Macmillan UK.

---

[12] The term Geordie refers to people from Newcastle and is often more closely associated with working-class men (Beal, 2009).

AMAND Maelle, 2019, *A sociophonetic analysis of Tyneside English in the DECTE corpus: The case of FACE, GOAT, PRICE and MOUTH* [online], Sciences du Langage, Université de Paris Diderot / Newcastle University, available at: https://u-paris.fr/theses/detail-dune-these/?id_these=4248

BAAYEN R. Harald, 2008, *Analyzing Linguistic Data: A Practical Introduction to Statistics using R*, Cambridge, Cambridge University Press.

BEAL Joan C., 2004, "English dialects in the North of England: phonology", *in* Edgar W. J. Schneider, Kate Burridge, Bernd Kortmann, Rajen Mesthrie & Clive Upton (eds.), *A Handbook of Varieties of English*, Berlin, Germany, Mouton de Gruyter.

BEAL Joan C., "Enregisterment, commodification, and historical context: 'Geordie' versus 'Sheffieldish'", *American Speech*, 2009, vol. 84(2), p. 138-156.

BEAL Joan C., CORRIGAN Karen P., MEARNS Adam & MOISL Herman L., 2014, "The Diachronic Electronic Corpus of Tyneside English: Annotation practices and dissemination strategies", *in* Jacques Durand, Ulrike Gut & Gjert Kristoffernson (eds.), *Handbook of Corpus Phonology*, Oxford, Oxford University Press.

BUCHSTALLER Isabelle, KRAUSE Anne, AUER Anja *et al.*, "Levelling across the life-span? Tracing the face vowel in panel data from the North East of England", *Journal of Sociolinguistics*. 2017, vol. 21(1), p. 3-33.

CHAMBERS Jack K., 1994, *Sociolinguistic theory: linguistic variation and its social significance*, Oxford, Blackwell.

CORRIGAN Karen P., 2012, "GOAT vowel variants in the Diachronic Electronic Corpus of Tyneside English (DECTE)", *in* Terttu Nevalainen & Elizabeth C. Traugott (eds.), *The Oxford Handbook of the History of English*, Oxford, Oxford University Press.

CORRIGAN Karen P., BUCHSTALLER Isabelle, MEARNS Adam & MOISL Herman L., 2012, *Diachronic Electronic Corpus of Tyneside English [Online]*, Newcastle University, disponible sur : https://research.ncl.ac.uk/decte (consulté le 10 décembre 2022)

CORRIGAN Karen P., MEARNS Adam & MOISL Herman L., 2014, "Feature-based versus aggregate analyses of the DECTE corpus: Phonological and morphological variability in Tyneside English", *in* Benedikt Szmrecsanyi & Bernhard Wälchli (eds.) *Aggregating Dialectology, Typology, and Register Analysis: Linguistic Variation in Text and Speech*, Berlin, De Gruyter.

COUPLAND Nikolas, 1980, "Style-shifting in a Cardiff work-setting", *Language in society*, vol. 9(1), p. 1-12.

DEUMERT Andrea, 1999, "Variation and Standardisation. The Case of Afrikaans (1880-1988)", University of Cape Town.

DI FRANCO Giovanni & MARRADI Alberto, 2013, "Factor analysis and principal component analysis", Milano, FrancoAngeli.

ECKERT Penelope, 2012, "Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation", *Annual Review of Anthropology*, vol. 41, p. 87-100.

ESCOFIER Brigitte & PAGÈS Jérôme, 1984, « L'analyse factorielle multiple », *Cahiers du Bureau Universitaire de Recherche Opérationnelle*, vol. 42, p. 3-68.

GENERAL REGISTER OFFICE, 1966, *Classification of Occupations 1966*, H.M. Stationery Office.

GHYSELEN Anne-Sophie & DE VOGELAER Gunter, 2018, "Seeking Systematicity in Variation: Theoretical and Methodological Considerations on the 'Variety' Concept", *Frontiers in Psychology*, vol. 9, article 385.

GRIES Stephan, 2009, *Statistics for linguistics with R. A Practical Introduction*, Berlin, Mouton De Gruyter.

GUY Gregory R., 2013, "The cognitive coherence of sociolects: How do speakers handle multiple sociolinguistic variables?", *Journal of Pragmatics*, vol. 52, p. 63-71.

HADDICAN Bill, FOULKES Paul, HUGHES Vincent *et al.*, "Interaction of social and linguistic constraints on two vowel changes in northern England", *Language Variation and Change*, 2013, vol. 25(3), p. 371-403.

HUSSON François, JOSSE Julie & PAGÈS Jérôme, 2010, "Principal component methods – hierarchical clustering – partitional clustering: Why would we need to choose for visualizing data?", *Technical report, Agro Campus Ouest*.

HUSSON François, LE DIEN Sébastien & PAGÈS Jérôme, 2001, "Which value can be granted to sensory profiles given by consumers? Methodology and results", *Food Quality and Preference*, vol. 12, p. 291-296.

HUSSON François, LÊ Sébastien & PAGÈS Jérôme, 2011, *Exploratory multivariate analysis by example using R*, Boca Raton, CRC Press.

JANSSENS Wim, 2008, *Marketing research with SPSS,* Pearson Education.

JONES-SARGENT Val, 1983a, *The TLS phonetic transcription scheme [Online],* available at: https://research.ncl.ac.uk/decte/transscheme.htm (accessed 5 November 2022)

JONES-SARGENT Val, 1983b, *Tyne bytes: A computerised sociolinguistic study of Tyneside*, Frankfurt am Main, Peter Lang.

LABOV William, 1972, "Some principles of linguistic methodology", *Language in society*, vol. 1(1), p. 97-120.

LABOV William *et al.*, 1981, "Field methods of the project on linguistic change and variation".

LAMENTOWICZ Mariusz, LAMENTOWICZ Łukasz, VAN DER KNAAP Willem O. *et al.*, 2010, "Contrasting species—environment relationships in communities of testate amoebae, bryophytes and vascular plants along the Fen–Bog gradient", *Microbial Ecology*, vol. 59, p. 499-510.

LÊ Sébastien, JOSSE Julie and HUSSON François, 2008, "FactoMineR: An R package for multivariate analysis", *Journal of Statistical Software*, vol. 25, p. 1-18.

LEBART Ludovic, MORINEAU Alain & PIRON Marie, 1995, *Statistique exploratoire multidimensionnelle*, Paris, Dunod.

MILROY James, 1996, "Variation in /ai/ in Northern British English, with comments on Canadian raising", *University of Pennsylvania working papers in linguistics*, vol. 3, p. 213-221.

MOISL Herman L., 2015, *Cluster analysis for corpus linguistics*, Berlin, De Gruyter.

MOISL Herman L. & MAGUIRE Warren N., 2008, "Identifying the main determinants of phonetic variation in the Newcastle electronic corpus of Tyneside English", *Journal of Quantitative Linguistics*, vol. 15, p. 46-69.

NERBONNE John & WIELING Martijn, 2017, "Statistics for Aggregate Variationist Analyses", *in* Charles Boberg, John Nerbonne & Dominic Watt (eds.), *The Handbook of Dialectology*, Wiley Online Library.

PELLOWE John, NIXON Graham, STRANG Barbara & MCNEANY Vince, 1972, "A dynamic modelling of linguistic variation: The urban (Tyneside) linguistic survey", *Lingua*, vol. 30, p. 1-30.

PICARDI Michela, REDAELLI Valentina, ANTONIOTTI Paola *et al.*, 2020, "Turning and sit-to-walk measures from the instrumented Timed Up and Go test return valid and responsive measures of dynamic balance in Parkinson's disease", *Clinical Biomechanics*, vol. 80, article 105177.

SOKAL Robert R, 1966, "Numerical Taxonomy", *Scientific American*, vol. 215, p. 106-116.

STRANG Barbara, 1968, "The Tyneside Linguistic Survey", *in* Ludwig Erich Schmitt (ed.) *Verhandlungen des Zweiten Internationalen Dialektologenkongresses*, Marburg, Franz Steiner Verlag.

STRANG Barbara, 1972, "TLS 015 Tyneside Linguistic Survey (summary)", Newcastle-upon-Tyne, Newcastle University.

STRANG Barbara & PELLOWE John, 1969, "The Tyneside Linguistic Survey: General handout on settings, aims, range and methods", Unpublished report, Newcastle University.

TRUDGILL Peter, 1972, "Sex, Covert Prestige and Linguistic Change in the Urban British English of Norwich", *Language in Society*, vol. 1, p. 179-195.

WATT Dominic, 1999, "Phonetic variation in two Tyneside vowels: Levelling in lockstep", *in* John J. Ohala, Yoko Hasegawa, Manjari Ohala, Daniel Granville & Ashley C. Bailey (eds.), *Proceedings of the 14th International Congress of Phonetic Sciences*, Berkley, University of California Press.

WATT Dominic & FOULKES Paul, 2017, "Tyneside English", *in* Raymond Hickey (ed.), *Listening to the Past: Audio Records of Accents of English*, Cambridge, Cambridge University Press. https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS1999/p14_intro.pdf

WELLS John C., 1982, *Accents of English: Volume 1*, Cambridge University Press.