Lexical Variation and Identity Practices within the Gypsy, Roma and Traveller Communities of the East Midlands

Peter Thomas Lee

A thesis submitted in partial fulfilment of the
requirements of Nottingham Trent University
for the degree of Doctor of Philosophy

October 2021

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank my director of studies, Professor Natalie Braber. You have been a guiding light and provider of support at each step of the way. I am sincerely grateful for the limitless encouragement and positivity that you have given me during the entirety of this research project. I would also like to thank my co-supervisors Doctor Laura Coffey-Glover and Doctor David Wright. You have all been incredible teachers and mentors and without whose support and direction, the development, progression, and completion of this work would not have been possible.

Other members of the department of Arts and Humanities at Nottingham Trent University have also been a great support. I would like to thank those who have been extremely supportive to me over the years I have been working on this project. I would like to thank Professor Martin O'Shaughnessy for his support throughout as interim assessor specifically, as well all members of the department who have given advice and encouragement throughout my journey.

I would like to extend my gratitude to those who have made the research possible and specifically those individuals who took part in the research for giving their support for this project. I would like to thank sincerely Mr Robert Dawson whose expertise, experience and enthusiasm was an enormous help to me in pointing me to individuals and resources that have been central to the development of this research. I would also like to thank the library staff at the East Midlands Oral History Archive, the staff at the Special Collections & Galleries at Leeds University and the staff at the Museum of English rural life at Reading University. I would also like to thank the volunteers at Leicestershire G.A.T.E who were instrumental in the advice and support they gave.

My appreciation, huge gratitude and special thank you go to my wife and daughters. You are all the best. You have given me everything I could need to stay on track and keep moving forward. This would not have been possible without your encouragement and the inspiration that you give me. I would also like to give thanks to my father for his help along the way. This was always appreciated and a motivation for this research.

Abstract

This thesis is the result of an ethical reflexive approach to the study of sociolinguistic lexical variation and identity of those who identity as Gypsy, Roma and Traveller (GRT) within the East Midlands of the UK. Whilst not the first study to look at lexis within these communities; lexical variation that exists between individual speakers has not been a central focus of the research. The study's focus is towards a variational account of lexis as a source of language variation, and, in doing so, addresses methodological approaches used to account for this variation between individuals. This variation is positioned within the Third Wave of sociolinguistics of the sociocultural interactional framework of Bucholtz and Hall (2005) who place identity as socially constructed through semiotic means. Previous research (Braber, 2015, p. 18) has suggested that the varieties within the East Midlands, minority groups more broadly, and the speakers of languages associated with Gypsy, Roma and Travellers as their communities of practice, have been overlooked in variationist literature within the study of linguistics.

An ethical reflexive approach has enabled observations of identity practices of those who identify as Gypsy, Roma or Traveller. This study highlights language practices of those that identify as Gypsy, Roma or Traveller and how lexis is utilised as an integral component of this practice. In addition to identity as social practice, I report on factors, which are influential in variation. I catalogued lexical variation and compared that with an original collation of data sets. These data sets have been used to establish factors influencing variation of lexis between speakers of the communities. These data sets were collated from a number of literary and historical sources, together with corpora, which have been established to collect data on variation within the Anglo-Romani dialects spoken within the UK. These findings give weight to the argument that lexis is used to perform identity alongside other linguistic forms. These practices relate to notions of self within discourse. This, together with the principle that lexical variation within these communities of practice, are rule based and governed by processes involving structured heterogeneity and are outlined within this thesis.

Chapter 1 Introduction

1.1 Introduction

The language and dialects, which are associated with the Gypsy, Roma and Traveller (GRT) communities within the UK have, thus far, been accounted for in terms of their linguistic typology. Studies have largely therefore only considered the linguistic features, such as: phonology and morpho-syntax of these dialects (Matras, 2003, 2010b, 2015, Hancock, 1984, Binchy, 2002, Ni Shuinear, 2002) with incidental commentary concerning the differences found between speakers, regional areas, and groups as communities that exist within a social network. In terms of dialect variation within the UK, traditional dialect research that examines regional variation has relied historically upon volunteer sampling that overlooks minority as well as marginalised groups (Orton and Dieth, 1974). Subsequently, subordinated groups and groups with a history of persecution have been overlooked. This has meant that questions relating to the influence the language and dialects of the Gypsy, Roma, and Traveller communities has had on regional variation, has not been investigated within the research and literature of dialect variation (Beale, 2015).

The present research will address this short-fall within the field of dialect and language variation research and literature of the British Isles by focusing on those individuals who self-identify as GRT by applying methodology for investigation of this minority group for the first time. The investigation focuses on the variation of non-standard forms reported by these community members by using a data collection method that collates lexical variation, thereby enabling comparability between informant responses (Kerswill, Llamas and Upton, 1999). The study also makes use of an ethical reflexive approach (outlined in 3.2) to account for variation used by individuals self-identifying as Gypsy, Roma or Traveller in the tradition of sociolinguistic research that considered language choice as a tool to construct social meaning though interaction (Bucholtz and Hall, 2005). This methodological approach has enabled data collection and analysis to focus on perspectives from within the community as interactions are participant led and between community groups. In addition, interaction itself is privileged as an area for contextualisation of the use and existence of variation where linguistic resources gain social meaning.

The Traveller, Roma and Gypsy communities are an integral part of UK culture and society and can be perceived as retaining their own culturally distinct life-style, beliefs (Hancock, 1992), which constitutes the specific focus of this study, language, and dialect (Hancock, 1984, Matras, 2010a). The extent to which these heritage varieties have crossed, mixed, as well as

influenced local dialects or maintained linguistic vitality is little understood and has been substantially overlooked (Beale, 2010). This study, with incentive to gain understanding of the largely neglected regional dialects within the East Midland districts (Braber, 2015, 2018), seeks to identify the linguistic and specifically lexical features of GRT Englishes found within wider social networks of this region. It also explores the variation that exists between speakers and the factors that influence this variation. Whilst the term 'GRT' represents a broad social categorisation, this study aims to explore this social category. Individual, self-ascribed identities are central to the exploration of identity as a factor for variation as social practice within this study. Individual perceptions of social network are also explored. Social Network theory has argued to have broadened (Diehl, 2018) to incorporate the idea of networks as cultural constructions. Networks are where narratives are shared, and symbols, endeavours and expectations of cultural practices (White, 2008) are reported by participants. By exploring identity and self-reported social networks as perceived by participants, this research will also highlight influences that these linguistics features have had on local variation. A focus of this investigation is to gain insight into the identity practices of those within the communities, and how these identities are constructed within linguistic interaction.

1.2 Theoretical Outline

This study of language variation is positioned within Penelope Eckert's notion of the Three Waves of variation studies of sociolinguistics and positioned within the Third Wave. The First Wave established correlations between broad macro-sociological categorisations (e.g. ethnicity) and linguistic variables. The Second Wave applied ethnographic methods to explore local categorisations, which constitute these broad social categories. More overt notions of linguistic variation as an identity practice such as covert prestige (Trudgill, 1972) have been explored within the literature. The Three Waves analogy originated from Eckert (2012). Eckert herself has refined the notion of the Three Waves as not mutually exclusive from one of the earliest studies of language variation, within the Martha's Vineyard study (Labov,1963) which she herself has reinterpreted as a Third Wave study. To this extent, the discussions of social meaning and discourse of agency, which is central to this thesis in terms of theoretical view, have always been a part of variation studies of this kind. A development within the study of language variation and identity is the idea that identity is not explicitly signposted through language choice. This has been defined by Silverstein (2003) as indexicality. The principle of indexicality is also central to this thesis (Ochs, 1992; Silverstein, 1976) which fundamentally explains the principles of how identities are constructed through social interaction. As such,

this research is placed within the three waves. As each wave has developed from the previous one, ideas have been made explicit from those that were previously implicit. This research however takes its influence from the principles which were embodied within the First, Second and Third Waves. This approach was outlined by Eckert (2018), and the principles from previous waves have therefore been embedded within the approach used here.

This study of language variation follows the theoretical premises established by Eckert's seminal account of adolescent language: *Jocks and Burnouts: Social Categories and Identity in the High School* (1989), a book-length ethnographic study that analyses the structures of social identity. This study examined adolescent linguistic variation and the social categorisation of Jocks and Burnouts. Whilst largely considered as central to a Second Wave, the implicitness of Third Wave notions was an important part of this work. Although not central to the analysis of this thesis, at the forefront of this seminal research was Eckert's notion of communities of practice formulated by Lave and Wenger (1991). Eckert used the concept of communities of practice to investigate the commonalities of language and its use as a means of constructing and co-constructing identity. Examples of Third Wave studies include those by Eckert that explored linguistic stye and the local construction of gender (Eckert, 1996), which was an example of Eckert's work in applying and developing Lave and Wenger's (1991) theory on communities of practice. Although the principles of communities of practice have not been applied in this study, the underlying focus on variation and identity as an interactional entity are a central consideration of the investigation. Other works that were founded on the principles of identity as being socially co-constructed, and for which this research is influenced by, include studies such as: Nerd Girls (Bucholtz, 1999), and on Lesbian groups (Jones, 2012). Within those studies adjustments of style, including associations between variance in meaning were established through observations of topic choice, and use of non-standard forms for example. Such observations reveal how linguistic devices as well as other semiotic symbols are used to negotiate power and identity within relationships. These notions are considered as a part of the principles found in Third Wave studies. This has been a major progression from earlier studies on linguistic variation (e.g., Labov, 1966; Lakoff, 1975). These concepts are central to Eckert and additionally Bucholtz' (2005) and were developed from original papers on performativity theory (Butler, 1988,1990,1993). The concept of performativity was based on the principle that identity and social meaning is an emergent feature specific to conditions of interactions. These principles were developed from the work of Austin (1962).

Bucholtz and Hall (2005) created a framework for the analysis of identity in a first paper of its kind on identity and interaction. Here, they outlined a set of principles that could be used for the analysis of identity produced within social interactions. The essence posited that identity was the product of linguistic and semiotic practice as opposed to an entity that was internal and psychological. Within that framework, identities were co-constructed through reference to macro-categories, which were created through a temporary interactional stance and based around emergent cultural roles that were created locally and between the interactional participants. In terms of indexicality (Silverstein, 1976) maintained that identities could be linguistically indexed through overt labelling, through implications being made, through adopting a particular stance on an issue or presupposition, through the use of style concerning linguistic structures and systems such as, choices of grammar, and, specifically for this study, through lexis. Identity was also established as based on the relationality of self in respect of others. This included notions of similarity and difference, ideas of authenticity and falsehood, and authority with the notion of delegitimising ideas within interactions. From this position, the notion of identity more broadly was complex and nuanced, and beliefs of self in relation to others were central to the interaction itself, as part of an interactional negotiation, subject to others' perceptions, self-representation and consisting of a larger ideological process and structure. To this end, the interactional data were analysed and understood according to those principles. The study explores lexical variation from an understanding that language is integral to identity. The approach, therefore, brings together principles determined in the three different waves, exploring lexical variation from the perspectives of individuals as speakers that self-identify as GRT. More specifically, the study seeks to address and answer the following three research questions.

1.3 Research Questions

RQ1: What constitutes non-standard lexical variation, with reference to semantic meaning and phonological features, for those who identify within a Gypsy, Roma and Traveller identity, within the East Midlands?

RQ2: How do individual background factors of a region: age, gender, attitude and identity influence non-standard lexical variation?

RQ3: What is the social meaning of non-standard lexical variation for those that self-ascribe to a Gypsy, Roma or Traveller identity?

1.4 Thesis outline

Chapter one of the thesis contextualises the research questions by giving an outline and review of the relevant literature. Chapter 2 explores language varieties of individuals that identify as having a GRT background and how it is of sociolinguistic importance. An explanation of the methodological advances, including the approach taken for this research are given in chapter 3. This chapter outlines the adaptations made to the Survey of Regional English (SuRE) method (Kerswill, Llamas and Upton, 1999). This constitutes a data collection method for language variation that has been adapted for this research project. A description and rationale for the fieldwork and approach that was adopted for this thesis is also detailed. Chapters 4, 5, 6 and 7 present an analysis and discussion of the lexical variation elicited through the SuRE data collection method. These chapters feature an analysis of identity as constituted in linguistic interaction. Chapter 8 concludes by synthesising the main findings from the data within the research questions, and by interpretating implications from those findings within the current literature and direction for future research.

Following the research questions and this outline comprises chapter 2. This chapter serves as a detailed account of the research context of those who identity as Gypsy, Roma and Traveller, which details the sociological and linguistic context for this investigation.

Chapter 2 Historical and Linguistic Background

2.0 Research context

To understand the motivations for this research and its questions, it is necessary to outline the sociolinguistic environment for this investigation. This chapter will therefore detail the context of the GRT communities in terms of historical position in section 2.1, including the linguistic context in section 2.2.

2.1 The Gypsy, Roma and Traveller communities

Historically, the GRT communities have been part of the UK's regional population, making up significant parts of the demographic of areas, such as the East Midlands. The official figure is 3,249 for the East Midlands based on results from Office for National Statistics survey (2011). The presence of a Roma ethnic group dates from at least as far back as the 16th century. Evidence for this can be found from a variety of official legal documentation and official correspondence. For example, from early court and municipal records an account of the imprisonment of 'about Christmas 1544, a number of Gypsies who had been imprisoned in Boston, in Lincolnshire, were by the King's command, shipped from there and landed in Norway'(Crofton, 1888: 11). Later accounts also identify regional incidents involving what many would argue to be early Roma émigré (Crofton, 1907; Winstedt, 1948) and are based on accounts found in official records from The National Archives and Acts of the Privy Council. Cressy (2016: 63) offers a more contemporary view on early historical accounts with similar references, such as an episode from Nottinghamshire in 1591 that documented people referred to as 'Gypsies' in relation to proceedings with the law. In this incident a number or 'bands of Gypsies' (Cressy, 2016; 63) were reportedly travelling to the then Gainsborough Fair and had been accused of 'pilferage' by local villagers. Documentation refers to the difficulties and perspectives of the authorities as opposed to the views and experiences of the subjects of the attention themselves. Not all accounts are negative however, as one incident recorded: one 'Gypsy' leader as standing out with a distinctive blue coat and even naming the individual as Thomas Jackson, which was an unusual practice in such documentation (Cressy, 2016).

The GRT communities, however, are not a homogenous group. The Irish, Scottish, English and Welsh Traveller groups are categorised within literature with often little or no Roma links although with cultural associations that may account for misconceptions and a lack of understanding. The Traveller population, in terms of documentation, historically separate accounts of indigenous itinerant groups predate records of Roma migratory groups. Crofton (1888) indicated how the surname 'Tinker' referred to once itinerant menders of pots and pans

and was traceable to at least as far back as the 13th century. Language also defines the Traveller population as a distinct cultural population. Linguistically, the Irish, Welsh and Scottish Travellers were documented with a legacy of 'Argots' or 'secret languages' referred to as Cant, Gammon and in academic contexts referred to as Shelta (Gmelch and Gmelch, 1976, Rieder, 2015). The extent to which the languages of these traditional communities have crossed, mixed, and influenced local dialects or maintained linguistic vitality has been little understood and substantially overlooked. This study, therefore, with an incentive to gain understanding of the largely neglected regional dialects within the East Midland districts (Braber, 2015), seeks to identify the linguistic features of Traveller and Roma English varieties in speech communities with a focus on regionality. The focus is to understand more fully how these features are used, have stabilised, and also the influence that these features have had on local variaties.

As a non-homogenous group, the question of cultural identity is worth exploring. The use of social categorisation labels such as the term 'Gypsy' is also of significant to research in that it examines language variation and identity. Issues of categorisation are part of an historic and ongoing discussion regarding the definition of itinerant communities, and those with a nomadic heritage (Ayla and Iris, 2016; Belton, 2005; Mayall, 1992; 1997; 2004; Okely, 1983; 2014;) with concepts that centre on ethnicity and of defining boundaries that distinguish one group from another. Hancock (1992), for example, has acknowledged that the term 'Gypsy' draws problematic definitions from those outsides of these communities, although Okely (2014) noted that community members self-ascribe the referent term 'Gypsy'.

The Traveller, Roma and Gypsy communities have retained their own culturally distinct lifestyles and beliefs (Hancock, 1992) whilst being an integral part of UK culture and society. The specific focus of linguistic variation within the community of speakers is considered as an integral characteristic (Hancock, 1984, Matras, 2010a). Questions arise as to what extent these communities, as self-defining within the British Isles and the East Midlands regions specifically, affiliate or sympathise with popular, historical, and academic accounts of Roma, Gypsy, and Irish, Scottish, English and Welsh Traveller representations. Areas that are a focus for exploration are the self-told stories and identity of these communities, linguistic variation, and attitudinal factors that promote or inhibit the use of dialect. This chapter first considers notions of commonality in terms of linguistic variation. It outlines the literature that has determined each community as broadly distinct in terms of an historical origin. Secondly, it introduces and outlines linguistic origins associated within these broad social categories, and how language features of these communities exhibit an integral position of a historical legacy

and heritage. This chapter describes Roma and Irish Traveller cultural origins, before moving on to a linguistic heritage associated with Roma and Irish Traveller heritage.

2.1.1 The Roma of the British Isles and Ireland

The Roma populations that travelled to the British Isles and Ireland are documented as migrating there during the early 15th century (Sampson, 1923, Crofton, 1888, Winstedt, 1913) and might have settled in Ireland as early as 1452 and subsequently in Scotland by 1460 (Crofton, 1888). Speculation of earlier movements to England by Roma groups from the then English-ruled French capital could also be given credence owing to the prospect of economic benefit and integration into the civil area of political hegemony. This, in addition to evidence outside of legislative accounts (Bataillard, 1884: 53) do confirm these migrations. In part, an expulsion from European territories would explain the more significant migratory movements of the Roma populations to more North-western European regions, since France, Germany, and Spain passed various and vigorous legislative acts between 1492 and 1504, aimed at outlawing Roma.

Within the UK, the migratory groups are now commonly referred to as Roma, Anglo-Roma, *Romnimos* alongside *Romani*, *Rom* and *Roma* or by themselves as *Romnical* and *Kalo* respectively (Hancock, 1984). Demographic data from the office for national statistics (2011) identified 58,000 individuals self-identifying as Gypsy and Irish Traveller. The census did not distinguish between those two ethnically-definable populations. This might be compared with an estimated 80,000 speakers of the creolised Anglo-Romani and a few hundred speakers of inflected *Romnimos* estimated by Hancock in 1984. These relatively recent figures have not included those estimates associated with more recent migrations of Roma populations from predominantly Eastern European territories.

2.1.2 Irish Traveller of the British Isles and Ireland

Less well-documented in terms of historical descent, Irish Traveller ethnicity has been viewed as a shared Irish heritage but with a commonly associated nomadic background and a population that has preserved a shared cultural, linguistic, and religious identity set apart from the sedentary or settled Irish population (Rieder, 2015). As with the Roma, Irish Traveller ethnicity has ethnic status with corresponding legal recognition which denotes their status as being protected against such actions that may cause threat by identification as a group or individual that results from inequality or prejudicial acts. This determines a legal status as well as determining Irish Travellers as a culturally definable group. It also acts as recognition that predates such legislation and affects the daily existence of those communities. Landmark cases

for ethnic discrimination date back as far as 2000 in the UK (Independent, 2000) and more recently enforceable under the Equality Act of 2010 (legislation.gov.uk, 2019). The situation was more recently developed in Ireland, and not until March 2017 (Oireachtas Report, 2017) was the ethnic status of Irish Traveller formally recognised by the Irish parliament.

The demographic population according to the national census of 2011, was 29,495 self-identifying as Irish Traveller within the Republic of Ireland (CSO, 2011), rising to 30,987 based on the status of usually residing in 2016 (CS0, 2016 (figures vary depending on the current status that is reported)). This would constitute 0.67% of the total state population. Of significance to ethnic divisions is the demographic distribution of age, with only three percent within the category of Irish Travellers aged 65 or over. This compares starkly with the 13.35 percent aged above 65 for the general population (Ibid). This an authoritative account of the relative distinction between qualities of life between the general populous and the subjugation of those who identify as Irish Traveller. In contrast to this record, the numbers for those residing in the U.K. are unclear as census data thus far do not differentiate by category for those who identify as 'Gypsy' and for those as 'Irish Traveller'. The amalgamated number is 58,000 for both 'Gypsy' and 'Traveller/Irish Traveller' identities in England and Wales (Office for National Statistics, 2011). Similarly, for Scotland the non-differentiated figure is recorded as 4,212 (Scotland's Census, 2011).

2.2 Languages of the Anglo-Roma and Irish Traveller communities

2.2.1 Romani Dialect within the UK

The language of the Romani populations is varied (Matras, 2015). In relation to the Anglo-Roma groups in the U.K., it is part of the cultural dynamic of the GRT groups and is part of a heritage descended through generations within the U.K dating back to the 15th Century. The variety found in the British Isles is maintained to a lesser or greater degree between the Anglo-Roma group members and its speakers. The familiarity and usage varies between families and between individuals as members of a social network of speakers (see section 3.1), and is evident in reports collected from the literature of the dialects of Anglo-Roma groups (Dawson, 2002, 2009, 2011) and wider ranging projects such as the Manchester Romani project (2021) which is a University of Manchester based project that seeks to collate dialects of Anglo-Romani found within the U.K. The Romani language possesses historically a cultural dimension that has given coherence to wider Roma populations, although this has not always been perceptible. Rüdiger's lecture in 1777 followed by its associated publication (1782), followed by a more widely published duplicate by Grellman (1783), which asserted that there existed Indo-Aryan

comparisons although contested by Sampson (1911) as the originator of that concept. This introduced and led to a more widely recognizable status as a distinct cultural ethnicity for Roma populations as part of a wider diaspora.

2.2.1.2 Early influences of Anglo-Romani

The work of Turner (1926) provided data that suggest earlier and prototypical language of Roma groups originated from a central Indo-Aryan group of languages. From this genesis, others (Woolner 1928; Beníšek, 2006) have attempted to define a split or migration away from central Indian populations as sometime during the 3rd Century BCE. Turner speculated from linguistic analysis how the Roma population migrated to North-West regions of India. This North-West region was originally thought by Turner (1924) to be the origin of the Romani language and therefore its cultural origin. However, this was later amended to show a more nuanced and complex source of origins. The language, Turner (1924) proposed, arose through contact between the migratory Roma population, with a pattern of borrowing over a period of a few hundred years with a North-Western Indian language contact. Matras (2002) pointed out the marginality of the evidence for the Central to North-Western split as the theory was based on vocabulary comparisons alone. However, Matras later wrote (2015) in support of an analysis of the Central group of Indo-Aryan languages as the spawning point for the Romani within the Indo-Aryan dialect clusters that were viewed as predecessors of modern Hindi, Urdu and Gujarati. The Romani language (Matras, 2015) is documented as showing a later split with central dialect clusters, which is revealed with borrowing and lending associated with the North-western Indian language contact and the Indo-Iranian and Dardic languages of Kashmir (Turner, 1926).

Language shift in the Indian subcontinent from Old Indo-Aryan (OIA) to Middle Indo-Aryan (MIA) in central regions corresponded with variations found within the dialects of Romani, providing evidence of contact. In addition to this, Turner (1926) provided linguistic clues showing a significant migration of the Roma population to the North-Western regions of the Indian subcontinent. This has been confirmed by way of Romani speakers, who retained many OIA variations, still maintained in the conservative North-West regions, but lost in innovative Central areas. More widespread changes that affected the whole of the Indian subcontinent did affect the Romani dialects (Turner, 1926) and therefore, with reference to the linguistic observations of Turner, demonstrated the evolutionary stages for an emerging Romani language, spawning in the Central Indian sub-continent areas and later migrating to a more isolated position in the North-Western regions of the Subcontinent.

Turner (1926) and more recently Matras (2015) highlighted parallels that revealed connections between central Indo-Aryan dialect innovations with characteristics of modern Romani dialects, and thus, provided evidence of contact. An example of innovations relating to OIA are the consonant clusters in the words of the central dialect of OIA *asman, tusme* 'we, you. pl.' that were reduced to 'amen, tumen', that are to be found in modern Romani dialects. This provided evidence of a connection between OIA and modern Romani dialects as the lexis are derivations in both meaning and form. These items also lost their sibilants in the language shift from OIA to MIA, which further connected modern day Romani with the central dialects of India and the Indo-Aryan languages that emerged there. Romani dialects did not follow all the changes that emerged between OIA and MIA associated with those central regions of India, and many features of OIA were retained within the Romani dialects. This retention of features of OIA aligned with modern Romani through an extended period in the North-Western region of India dating back to an early Medieval period where such features were to be found. Matras (2015) pointed to a period midway in the first Millenia where such features were retained, whilst others began to mutate . One example of the retention of OIA features was the intervocalic dentals. An example is in the modern Romani word '*gelo*' /dʒelɒu/ meaning 'gone' which can be compared with '*gata*'/dʒætæ/ in OIA in terms of meaning and phonological relationship. These retained features have either been lost or simplified in modern Central languages, as observed in Turner's (1926) work, which indicated Romani regional origins. Changes that influenced all Indo-Aryan languages such as the simplification of consonant clusters can be seen to affect Romani, which is evidenced in modern dialect. The OIA word for 'snake' *sarpa* has been reduced to *sappa* for MIA, and OIA 'blood' *rakta* becomes *ratta* in MIA or *rat* in its modern Romani equivalent (Turner, 1926). Table 2.1 illustrates the retainment of OIA features outlined above.

Table 2.1 The reduction of OIA to MIA, to Modern usage

English	OIA	MIA	Modern
blood	*rakta*	*ratta*	*rat*
gone	*gata*		*gelo*

Such parallels have provided the strongest evidence that has enabled a temporal line to be drawn, tracing a population with a path of movement for the speakers of the Romani language and dialects. Romani today is classed as a New Indo-Aryan dialect with Old and Middle Indo-

Aryan influences. Figure 2.1 highlights the distribution of New Indo-Aryan languages across India. The Romani dialect split is speculated to have occurred around the turn of the first Millennium.

Figure 2. 1 Map showing modern distributions of New Indo-Aryan Language (NIA) Kolichala (2015) in Hock (2016)

See unabridegd version

2.2.1.3 Migratory influence upon Anglo-Romani

The Romani language spread and this is revealed in the influence of language features outlined by the literature that identified language contact features (Matras, 2002, 2015). Linguistic parallels with the languages of those regions have shown language contact from the development of the proto-Romani within the North-Western regions of India and a move to linguistic developments as language contact migrated to neighbouring areas. The migratory route is illustrated in Figure 2.2 and shows a possible route from the ranges of the Pamir which adjoin the Himalayas with Afghanistan, moving on to associated influences of Iranian and Caucasus languages with their associated geographical locations. Romani words such as *grast* for *horse* and *bov* for 'oven' are of Armenian origin, and *baxt* 'luck' and *tover* 'axe' are of Iranian decent. Pott (1844-5) and later Sampson (1923) suggest a continuation of movement, whilst Matras (2015) indicated concurrent borrowing and assimilations as a result of a

multilingual settings could just as easily have been the cause. The peripatetic and itinerant characteristics are evidenced through linguistic borrowings. The extent to which the then Romani populations had simultaneous language contact between Armenian and Iranian speakers was certainly possible given geographical locations and distributions of those population suggest a multilingual society, a theory which has been put forward by Matras (2015). Evidence of language contact between Armenian and the Iranian languages has promoted the plausibility of an extended multilingual period of contact between these varieties. Variations that existed within the Anglo-Romani communities demonstrated not only linguistic features of those languages, but also from language contact from areas passed through as has also been illustrated in Figure 2.2.

Figure 2.2 Map speculating Romani migratory route from India. The map shows modern regional divisions (Taylor, 2014).

See unabridged version

2.2.1.4 Anglo-Romani and Greek

The Romani speaking populations are shown to have lengthy contact with the Byzantine Greek empire. This is illustrated in Figure 2.3, which shows a migratory route during the mid to late period of the first millennia. Evidence for this contact again arises from cross-analysis of linguistic data, which was examined, principally, by Turner in his seminal analysis and most notably re-evaluated by Matras (2002, 2015). There are many examples of earlier Greek borrowings with examples such as the Greek word *drom* 'way' or 'road', which are integrated into Romani dialects, and seen in modern usage. The plural form of *drom* in Romani dialects however, receiving the Romani inflection 'à' as in '*droma-à*'. Similarly, the Greek word for 'week' *kurkò* receives the plural marker *kurkè*, which follows the pattern for Romani words as seen in Romani 'boy' *chavò* and 'boys' *chavè*. Later integrations kept much of the original

inflections with many words incorporated into Romani maintaining their Greek superstratum grammar when investigating the language in modern usage. Significantly, for modern Romani varieties, a large amount of vocabulary was introduced during that period, with speculations relating to an extended period of contact between the Romani population and the Byzantine Greek speaking Empire. The influence of Greek can be viewed as a reflection of necessity for trade within the Greek empire of the time rather than evidence of location or of movement. The inheritance of Greek within the Romani dialects might therefore have been a result of language contact and fed by economic opportunities from this contact with the Byzantines.

Matras (2002, 2015) and Elsik and Matras (2006) emphasised the significance of the Greek influence upon the dialects of the Romani language tree. They estimated approximately 200 words were from Greek in any one dialect and approximately 1000 lexical word roots could comprise the total. This means Greek is one of the larger and more significant linguistic influences upon the Romani. In addition to vocabulary, the period of language contact with the Byzantine Greek era Europe manifested an inflectional system carried on to the present. Matras (2015) has identified the example of 'doctor' a modern borrowing into Romani but inflected for nominative case using the Greek nominative ending '*-is -o,* or *-os'* as in Romani *doctoris, doctor, doctoros*. The evolution of the Romani language to a typologically European language with verb-object word syntax, from a contrastive object-verb syntax of Indi-Aryan languages began to emerge during this period of contact. The language adopted many of the grammatical characteristics of Greek and Balkan languages such as a structural distinction between factual or epistemic clauses (*I think that*), including intentional or volitional clauses, and non-factual clauses (*I wish that*) (Matras, 2015). The characteristics of an Indo-Aryan language with European influences is revealed through these analyses. The stages from a prototype to an Early Romani language period can be shown through these comparative analyses. These linguistic footprints have allowed for speculation involving migration over several hundred years from the Indian subcontinent through to European territory and into a broad dispersion and on to a diasporic populous as seen today.

Figure 2.3 Map showing migratory movement to (mid-first millennia) and from (18th century) Byzantine Empire based on 6th Century representation (DeCarlo, 2017)

See abridged version

2.2.1.5 The transition to Anglo-Romani as dialect: typologies

The dialect formation of Romani is rooted in the premise that Romani up until the period of the collapse of the Byzantine Empire was a single proto-Romani spoken by a coherent population. This model was composed by Miklosich (1872-1880) to illustrate how the various branches of dialects resulted from a decentralisation from a single or multiple proto-type and root of Romani, assumed to be during or at around the time of the end of the Greek Byzantine period or during the middle of the 14th Century. The quest for historical places of origin for a speech community could assume only that those speakers of a proto-Romani dialect held a period of language contact with the Greek Byzantines, as would be expected given the geopolitical circumstances of this period.

The coalescence of speech communities and the influence of Indo-Aryan languages played a significant part in the formation of the proto-variety. Miklosich's model illustrated the transference and splits between the speakers, leading to the subsequent results for the dialects spoken by each speech community. The widely accepted classification system for the dialects of Romani is based in terms of branches set out in this model and represents an example of an isogloss associated with this model is illustrated in Figure 2.4 below. The map illustrates part of a classification system developed to distinguish dialects of Romani in terms of their regional distribution patterns. The verb endings are displayed within regional isoglosses on the map and determine or suggest an influence of geographical distributions of Romani dialects.

Figure 2.4 Map showing an example of Mikloshich's modelling illustrating loan verb adaptation markers (*The Classification of Romani Dialects: A Geographical-Historical Perspective* (Matras, 2005)).

See unabridged version

Matras (2015) made a distinction between the settled groups of Roma and those who were nomadic during the period of migration out of the vestiges and subsequent reformation of governance in the area of Asia Minor. The principle that Romani are nomadic is seen in this light as part of a pattern of movement associated with compulsion, and heavily dependent on the context given by the dominant authorities and administrations at any given time.

The period that followed represents the movement of proto-Romani speech communities into European territories, which continued for centuries and until the present day in the case of the British Romani population. The recording of anti-Gypsy legislation dominated the documentation of this extensive period and examples of which were referenced previously (section 2.1) and dominated many of the accounts of the presence of Romani communities as distinct in their culture. The first account of the Romani language in the British Isles is given in Boorde's list written in 1547. This account revealed a picture of dialects differentiable in much the same way as can be seen today, when compared with other recordings of Romani dialects (van Ewsum (*c*. 1560; see Kluyyer, 1910), Vulcanius (1597; see Miklosich 1872–80), Evliya (1668; see Friedman and Dankoff, 1991) and Ludolf (1691; see Kluge 1901). Some authorities (Boretzky and Igla 2004; Boretzky, 2007) cited this phenomenon, signalling the formation of dialect distinctions before the European integration period of the last few centuries, and one which took place prior to the extended Byzantine contact period.

Others (Matras, 2002; 2015) refuted the possibility of large and organised parties travelling together and settling in specific regional territories and subsequently establishing speech communities associated with specific local areas. Instead, evidence has been assembled to align the influences of co-territorial languages. Matras (2015) has listed vocabulary, prosody, semantic innovations and effects upon substrate (Romani) morphosyntax with much of the processes involved in innovations that have taken place when much of the interest in Romani language and culture began during the 19th Century. The paradigm that Romani spread from a single proto-typical form and then metamorphosised into distinct regions seems to contradict the documentation of the earliest recordings during the early European period of Romani migration. As approximately during the middle of the 16th Century Romani dialects are documented as being already distinguishable, however. The linguistic influence of local and regional language of host European populations upon dialect formation does not rule out the plausibility of dialect variation prior to this movement of Romani speakers. The coalescence of dialect within each principality or region can, therefore, be seen as an emerging feature of a population that shared linguistic and cultural backgrounds. Nonetheless, evidence for dialect

variation in an early linguistic form of Romani has existed in the form of analysis of geographic distributions of linguistic variables as depicted in this work.

Using a comparative method of analysis concerning modern variations, random occurrences as opposed to geographical patterns of sets of associated features related with Early Romani can be viewed as signals of variation prior to pre-European settlement, and this was present amongst many of the dialects of Romani using this method as shown above, and would be expected within a study of variation, such as this, and found during this research. Variation has been conceived as pre-existing and predicted as uninfluenced by local triggers (Matras, 2002, 2015). This model has posited tangible evidence for the pre-existence of dialect variation although limited in its predictive capacity to estimate the extent of the variability. This would assume that the simplification processes and levelling that acted upon Early Romani inheritance could also likely have occurred as a result of a language that emerged through contact. Whether the inheritance of Indo-Aryan language and geographical and regional traces are correlated is questionable although commonly assumed and one for further investigation.

2.2.1.6 Modern Variation

The predisposition for local and regional changes within the dialects of Romani are therefore considerable given the prospect of variability prior to movement into European regions. Coupled with patterns of assimilation, this has led to many dialects within specific European territories and across the European continent. Hancock (1984) quoted at least 60 dialects of Romani as extant (Hancock, 1975, Wentzel and Cherenkov, 1976, Kaufman, 1979). Whilst the distinctions between dialects have been a centre for earlier work on variation, whose varieties are referred to under the term Anglo-Romani, there has been relatively few if any that have focused on variation as a socially interactional phenomenon and research that viewed language variation as a sociocultural phenomenon, as opposed to an external entity and definable as a meta-interactional artifact.

In terms of dialect branches, the dialects have been divided ostensibly into five explicit regions (Bakker & Matras, 1997, Elšík & Matras 2006). The British Isles or Anglo-Romani dialect is seen, in this distribution, as part of the Northern dialect group which itself is further categorised into East (German, Scandinavian) and West (Polish-Baltic) and shown in Figure 2.5.

Figure 2.5 Map of Romani Dialect distribution showing conventional location of Romani dialects (Matras, 2010b)

See anabridged version

The British Romani dialect is seen in this categorisation as a distinct sub-category of the Northern western type. Iberian, Central, Vlax and Balkan are the four other typological distinct branches that make up five attributable groups, albeit a 'rough' categorisation (Matras, 2010b: 28). Table 2.2 shows examples of features that can be used as a diagnostic for what have been labelled as dialect branches that are representative of regions in the Central and Vlax regions. The table illustrates a pattern of distribution of language features such as suffix or lexis that categorise a dialect as a Vlax dialect of Northern or Southern for example.

Table 2.2 Example diagnostic features used for dialect branches (Matras, 2013)

See unabridged version

Romani spoken in the British Isles or Anglo-Romani has been documented sporadically and largely ignored up until the 18th Century. As for the speakers, Hancock (1984) described two

distinct waves of migration for the community of speakers, which he labelled as *Romničal, referring to* the initial migration, and the second wave after an abolition of slavery in the Balkan states during the period after 1864. Hancock continued by suggesting that those two populations possessed little in common although clearly a linguistic and shared historical heritage affected and attested to a shared sense of culture and communal past. For Crofton and Smart (1875: 219-21) the dialect spoken by the Anglo-Romani or British Romani, were composed of two comparable 'Old' and 'New' dialects, with a loss of verb inflections, Romani prepositions, inflected personal and possessive pronouns, genitives and subordination replaced by English in the 'New' variety. Lexis constituted the main constituent of the modern variety with fossilized personal pronoun forms, demonstratives and negation particles, including the rare use of inflections (Winstedt, 1948:103). Matras (2010, 2015) attributed this shift as moving from a matrix language, defined as a language used to frame, contextualise and exchange information, towards the use of a code that demonstrated solidarity, and symbolised kinship and for a variety of socio-pragmatic special effects. The movement from a communication tool to a system for emblematic display has been, according to Matras (2015), past the point for what he termed as a 'functional turnover' (Matras, 2015: 39) for this dialect of the Romani language. This suggests a significant distinction in the changing of use of Anglo-Romani as a community language.

2.2.2 Irish Traveller Language and Dialect

The language associated with Irish Traveller ethnicity possesses several referential terms and has often been referred to in academic contexts as Shelta, first used by Leland (1880). Binchy (2002) regarded the use of the word Shelta as a sign of respect accorded to the speakers of the varieties that this term covered. She referred to both Gammon and Cant as terms that the speakers of these languages used as reference terms for their own language, and the use of the term Shelta as an academic term used by those outside of the speaking communities, and one which set a boundary between a language for academic discussion and a language as referred to by and for its users. The term Shelta dates back to the Victorian Gypsy-folklore. The first mention of the term Shelta was by Leland who published a chapter '*Shelta, the Tinkers' Talk'* in *The gypsies* (1882) in the second half of the 19th Century. The term Shelta had been retrospectively etymologised most controversially by Macalister (1937), a well-known lexicographer for Shelta varieties. Macalister hypothesised an elaborate derivation that led to the coining of the term Shelta from the Irish word *bèarla,* an old Irish term for 'language'. However, Ni Shuinèar (2002) proposed a simpler association with the Irish word *siùltòir*

meaning ‘a walker’ which itself was a variation on the word *siùil* ‘to walk’. Ni Shuinèar (2002) assumed that the final ‘r’ was lost through Anglicisation and hence forth the term Shelta became a reference for the variety of language spoken amongst the Irish Traveller communities, and notably not a term used by any of its speakers. More recent studies of the varieties spoken by Irish Traveller populations highlighted that amongst the speakers of this in-group language and dialect, the term Shelta is not known (Binchy 1994; Cash, 1977: Hancock 1984).

Nì Shuinèar (2002), in addition to deriding the insistent use early researchers employed of the term Shelta, critiqued the lack of depth of previous linguistic studies. The critique highlighted those that studied Irish Traveller language or early Lorests, and focused on a synthesis of R.A. Stewart Macalister’s 1937 treatment in ‘The Secret Language of the Irish’ that investigated Shelta. In this early account there was an examination of vocabulary with a tendency to treat its subject as an exoticism rather than an analysis of a language, its composition and its use. Nì Shuinèar further condemned how initial studies of the varieties spoken by Irish Traveller groups have been viewed, not as a language but as a supplementary jargon that had developed through modification of Irish Gaelic, following Leland’s (1882) work. This assumption has been long standing and based on presuppositions of Leland’s, that Shelta varieties were based on or formed out of another language, rather than through patterns related to contact, maintenance, and founded on a distinct linguistic form.

2.2.2.1 Origins of a Language

Opinions from within the Traveller communities have shared a suspicion of scholastic interpretations of the origins of ‘the Cant’ and ‘the Gammon’ (Walters, 2002). Macalister (1937) supposed that forms of Shelta were derived from modern Irish and proposed a systematic analysis of the derivational process. However, Macalister’s scholastic rigour has been scrutinised. Although a significant recording of lexical items was one of his most notable contributions to Irish Traveller linguistic studies, more recent observations revealed limited validity for the data and the participants involved in its collection, including any subsequent interpretations that emerged. Macalister assumed a thorough cataloguing and recording of contextualised use but given the lack of recording and transcription options of the time with an emphasis on lexis, many of his conclusions were tentative at best. The linguistic processes of transformations are seen as substantially based on intuitions, owing to inadequacies inherent in the data collection. The principal conclusion was that Shelta developed from Irish was largely based on modern observations of a comparable word order, inflections, and use of relatively few articles. The system of derivation worked out for Irish Gaelic to develop into Shelta has

also been criticised as being too complex to the verge of being universal for any language's reconfiguration from Gaelic to Shelta. Discussions proposing earlier origins have also been widely debated. Older words for 'priest' and 'death' are used by Travellers (Binchy, 1994) and purportedly found on 6th Century Pictish tombstones (Walters, 2002) and suggested older influences or earlier sources of origin. The question as to whether Shelta was an ancient language or a more modern development that was related to a socioeconomic status often assumed that the one negated the other.

Sampson in 1886 (Liverpool University special collection) and Meyer (1891, 1909) began an academic collection and analysis of the Shelta language and proposed an ancient origin, later to be contested by Macalister. Their observations of an earlier inheritance were based on an archaic term use found in old scriptures. Methods for disguising words were found in old monastic texts and evidence for lenition, the weakening or softening of articulation placements, of older forms of Irish were also uncovered. The influence of Irish, although not isolatable as a source of origin in terms of linguistic analysis, was shown and contact with the language has had an influence on the development of Shelta. Binchy (1994) made clear, however, that Irish and Shelta were two distinct codes; with maintenance and use founded on ethnicity and heritage rather than on transactional and perfunctory use of a secret code or argot. Her point is illustrated by the continued use of Shelta amongst those ancestors travelling from Ireland post American emigration circa. 1848-1850. With Irish, in addition to Shelta being spoken, the need for Shelta as a secret code became unnecessary, yet as Binchy argued, the language has been well-maintained amongst those Shelta speaking émigrés.

2.2.2.2 Conceptions of a Language: Ancient or Modern

Two camps have been established with one represented by Binchy (1994) where the less prestigious and fragmented speech community has been eroded and subsumed, and a need for an independent grammar, and ethnicity was symbolised through lexis alone. By way of contrast, Ò Baoill (1994) also focused on the syntax of Shelta but without discussion of ethnicity and origins. From those observations a significant sharing of English lexis was highlighted together with a large number of non-English words, with an identical grammar and a language used for specific occasions, with a well-defined purpose, although no register differences were identified. Ò Baoill (1994) speculated that the language must have developed amongst bilingual speakers of Irish and English at a time when the two languages would have had significant contact. This concept, however, ignored the era for the forming of the language, was illusory

as it did not satisfy an understanding of why it emerged amongst some particular speech communities and not others.

2.2.2.3 Influence of the Irish Language: Word Play and Smithery

There was no dispute however, in the observation that English and Irish have had a significant influence on the Shelta language and its dialects. Ò hAodha (2002) considered the influence Irish has had on the lexicon of Shelta through data from a 'West of Ireland Travelling woman' (2002, 47). The data illustrated how Irish has influenced the formation of the Shelta lexicon and highlighted influences of Irish in the use of Shelta. Of note was how speakers of modern Shelta have limited or no Irish Gaelic proficiency although it had been passed down in its Shelta form. This assumed that any Shelta speaker with influences of Gaelic in modern Shelta usage had partly inherited this from a bilingual past. In respect of the Irish influence, Ò hAodha (2002) pointed out that a study of the large number of non-Irish lexis descent relating to Shelta has yet to be carried out, however.

Theories of the exact circumstances of origin have differed. However, Irish and Gaelic influences were clear (Meyer, 1891; Hancock, 1984; Grant, 1994; Binchy, 1994; Ò Baoill, 1994). The integration or borrowing from Irish as well as English followed a series of patterns, as outlined by MacCalister (1937), and conformed to a less complex set of rules used by other minority groups associated with an argot such as the nomadic Bards, who, in Celtic culture, were professional story tellers and poets. Macalister (1937) with this development, suggested a theory of assistance or 'literate help'. Here, the language was seen as developing through patterns of borrowing and integrations from scholastic and monastic influencers who were seen as developing a system of word play such as transpositions and reversals. In this sense, although seen largely as a language with a non-literate history this was not the perspective of all. The poetic and innovative characteristics of the language were observed as a feature of a language with a literate past and one closely integrated with the narratives of the nomadic Bards (McDonagh, 1994). The connection between surnames of Ward, and the Irish *Mac an Bhàird* linked many Irish Travellers as a 'son of the bard' and therefore, their language with the bards and most notably their associated wordplay and word-smithery. The language and the dialects of Shelta, from this perspective, were seen as being maintained, inheriting a characteristic of innovation in addition to the older lexical heritage, associated with the documentation of poetry and characterful word play as part of a cultural inheritance.

2.2.2.4 Modern Variation

Within the dialects of Shelta there were arguably significant differences between their form, usage and pronunciation. Acton (1994) pointed to variations between groups and families of speakers of Shelta. He cited an observable distinction between the varieties of Southern Irish, those of Ulster, the Northern Scottish, the Romani influenced mid-Scottish, and the South Welsh dialects. These varieties were distinguishable in terms of lexis and their pronunciation and were distinguishable between speakers of these varieties of Shelta.

In common however, were patterns of word play and Meyer (1891) summarised the significant few rules of transformation to four. The first rule was the reversal of mostly Irish words, a second being prefixing an arbitrary letter(s) to a word, a third was substituting an initial letter(s) of a word, and a fourth transposing letter(s) or formation of an anagram. Substitutions were the most frequent method of word formation, with reversal or 'back-slang' a close second. The reversal method of word play was associated with older forms of Shelta by older generations of speakers (Ò hAodha, 2002). One other word formation method was metathesis as in /æks/ for ask, and embodied a process that was no longer developing and was only observed for Irish words and for the transposition of vocabulary.

2.3 Context of Research

Whilst previous research on languages of GRT communities has highlighted the linguistic characteristics as discussed within this chapter, few have done much to highlight the characteristics of language use as performative within a sociocultural model view as used by those self-identifying within these communities. The positioning of identity within social contexts has been central to notions of the role of language as an interactional tool (Bucholtz, 1999) and one that has been arguably underexplored in the context of minority ethnic identity. Matras has outlined the use of dialect and language used by a participant pool of forty individuals from various regions within England who self-identified as English or Welsh Gypsies and who identified as Roma or *Romanichals* (2010a:132) and speakers of Anglo-Romani. The dialect observed has been presented as an in-group vocabulary used amongst those with a common socioeconomic profile that 'serves as a kind of formalisation or conventionalisation of a mode of communication that calls on the hearer in a rather explicit way to activate a specialised presuppositional domain when processing an utterance' (Matras, 2010a:132). In this instance, speakers used an in-group lexicon referred to as 'Anglo-romani' (Matras, 2010a:1) to signal group solidarity (Matras, 2010a), and to reflect a group's attitude towards an actor, event, or other referent. Binchy (1993; 1994; 2002) discussed the in-group

lexicon of Irish Traveller communities in her discussion of Shelta. Her studies again examined direct observation of participants and their language from within the Traveller communities. She observed that the Travellers she studied shared a common background knowledge and could be classified as a closed network group. Binchy wrote that the use of Shelta was used when a high degree of contextual understanding was required to follow meaning. Domains of use included making a living amongst non-Travellers, maintaining boundaries between non-Travellers, and use within Traveller-community members for intimate subjects, for example. Thus, the language functioned as a code for maintaining a close community, kinship and social code. These studies highlighted the use of language amongst the GRT communities as a prime source of information about the role of language in social practice.

Historically, studies have focused on lexical variation within Romani dialects found within the British Isles. For example, Romani language can be traced back to an early transcription of cited Romani sentences documented by Andrew Border in 1547 (Boorde, 1547). A comparably dated list of words also cited as Romani lexicon has been detailed by Bakker (2002). Significantly, the words were purportedly inserted into English syntax and not Indic syntax as outlined within the contextual outline here, and will be discussed later within the examples from the analysis of this thesis. It appeared that this code, which has been classified as an inserted lexicon (Matras, 2010a), and far from disappearing or dying might have thrived in its present form for a much longer period than would have been expected given the pattern of decay and language death of minority dialects and languages. The nature concerning the longevity and existence of these dialects and languages of these ethnic groups has been a central focus of this investigation. The present study investigates language use and specifically lexical variation and its use as a social practice. The next chapter outlines the methodological procedure and rationale for the selection of research tools used to collect lexical, attitudinal and background data for the purposes of answering the research questions set out in section 1.3 and reiterated within the methodology chapter.

Chapter 3 Methodology

3.1 Introduction

This chapter outlines the methods used to collect and analyse the data in this thesis. The study combines the use of empirical data collection using adaptation of the SuRE methodology (Llamas, 1999). This incorporates quantitative attitudinal and identity measures and qualitative ethical reflexive methods by using semi-structured interviews. The purpose of using qualitative measures was to ascertain the social meaning of variation and maintenance of language and dialect of individuals. The linguistic practices were considered relative to a broader concept of social networks, of which participants identified as being members. Linguistic practices, in this sense, were formed through common interests or positions, and provided s a link between an individual, group, and place. The notion of 'community' is based around self-reported identity. In this sense the 'community' is an exploration of common endeavours and shared experience and is based on self-perception.

The data collected from the semi-structured interviews are positioned alongside quantification of attitudinal and identity measurements. Previous research carried out by Kerswill, Llamas and Upton (1999) has used the same data collection methods developed by Llamas (1999). These SuRE methods outlined within this chapter, have shown that use of non-standard lexis is associated with measures of attitude and identity in addition to social networks (Milroy, 1987). Social Network here refers to a broad network of language users. The network differs in the density. This is measured by the numbers of links there are between members of a network. It also differs in how central or close each individual is to one another, and how many social ties each person has to the next person within the network. For the purpose of this study, it refers specifically to a participant's rudimentary perception of local identity of those they

considered ethnically related and as part of their community. This concept was based on a self-reported measure (see section 3.8.1 for further discussion).

The SuRE research methods have not been used to collect data for the analysis of mixed language varieties. Mixed languages can be defined as two source languages that are combined as with cases of community bilingualism (Meakins, 2013: 159) and relevant to the languages of Anglo-Romani and the Irish Pavee Cant discussed in chapters 1 and 2. Identification of an individual's lexical variety is central to this thesis and one which the individual associates with their self-ascribed ethnic identity. The lexical variation is considered as part of the participants' 'mixing' of language. As such, this is considered conceptually as a lesser known variety of English. Historically the mixed language of Anglo-Romani and Pavee Cant has been labelled within the dialects of the GRT communities. This study seeks to look in depth at the individual practices of speakers and specifically in relation to their own self-described attitudes and perceptions of identity. It is argued that there is a gap within the literature and field of variation and varieties within the U.K. with many studies identifying typological features (Schreier, et al., 2010) whereas this study looks at an emic (insiders) description of identity and language practice. The value of insights gained from the current research and its analyses are a result of the interpretations made possible from the ethical reflexive approach to interview and fieldwork data collection and its analyses.

The use of the semi-structured interview provides qualitative data to explore individuals' own account of language use in practice. This method enables an exploration of the social meaning of lexical variation whereby participant evaluations and judgements of language description and their use can be discussed and considered as part of their own experience. This forms a robust tool for representation of variation in practice, allowing for validity in terms of what they represent and reliability in relation to comparison with others findings. These methods are selected as they are specifically tailored towards answering the three questions of the research questions set out in chapter one. These are:

RQ1: What constitutes non-standard lexical variation, with reference to semantic meaning and phonological features, for those who identify within a Gypsy, Roma and Traveller identity, within the East Midlands?

RQ2: How do individual background factors of a region, age, gender, attitude and identity influence non-standard lexical variation?

RQ3: What is the social meaning of non-standard lexical variation for those that self-ascribe to a Gypsy, Roma or Traveller identity?

These questions relate to social and cultural factors that are associated with variation and use of dialect that are relevant to individuals' ethnic and cultural identity. The first question concerns lexical variation and specifically the mixed languages relevant to Anglo-Romani and Irish Traveller heritage. Part of the enquiry looks at regionality and attitudinal factors that relate to its use. Quantitative measures that consider self-descriptions of identity have been integrated into the methodology to assist the determination of influence for the linguistic variation observed. The self-reported identity measures show an individual's self-ascribed identity (Table 4.3). For the semi-structured interviews, discussion centred around language and a self-perception of its use within the speakers own social network use. The combination of these two methods enables both collection of ethically reflexive data central to the analysis of this research, and attitudinal and identity information that go alongside discursive accounts of cultural experience from the researcher (Llamas, 1999). This approach was developed in consideration of Eckert's (2000) work whose seminal study combines a fine-grained approach with a methodologically robust analysis of variation within micro-communities. The study applies an ethical reflexive approach that uses semi-structured interviews as part of the methodology in addition to attitudinal measures analysed within chapters 4 through 7. Outlined here are the methods which will contextualise the qualitative methodology from the SuRE methods used and adapted in order to ensure rigour and reproducibility of project. Section 3.7 outlines the method and theoretical principles for collecting and analysing the quantitative data. The quantitative data is subsequently utilised within chapters 4 to 7.

I start by introducing an ethical reflexive approach in section 3.2. I will also focus on how ethical and reflexive awareness has been used in previous studies within the field of linguistic research and specifically on minority and regional language and dialect variation. I will contextualise this approach with an outline of previous research done in this area in section 3.2.1. I then move on to discuss my own experience of conducting ethical reflexive research in section 3.2.2. Here, I discuss my own experience of conducting research and my background in relation to the focus of the study. This is followed by an outline of sampling methods used in 3.3. The participants and their background are described in detail in section 3.3.1 and the contexts of data elicitation are described and outlined in Table 3.0. The collection of data and how this was managed and adapted from Llamas' SuRE method (1999) is detailed in section 3.4. The method for lexical elicitation is provided in 3.5 and how this was piloted is detailed in

section 3.5.1. The description of the lexical elicitation tool is then described in detail in section 3.5.2. The semi-structured interview is then outlined in section 3.6. The quantitative measures of attitude and identity relation described in section 3.7 with rationale for the purposes of this research outlined. This is followed by a description of theoretical framework applied for the analysis of meaning within interactional dialogue in section 3.8. The study's ethical considerations and foundations are explained in detail in 3.81. This is followed by limitations (section 3.8.2) relative to implementation, discussion and conclusions that follow.

3.2 Ethical Reflexive Method

The beginnings of a reflexive approach to sociolinguistics methods can be found in Hymes (1962). His first studies were on American Indian cultural and linguistic diversity as an approach that can be viewed as an answer to a Chomskyan quest for universals. The approach by Hymes can be positioned as an opposing argument to Chomsky's and argues for a process that widens the linguistic scope of variation study by describing diversity. This counters the quest for notions of correct use of language and the idea of ideal speaker judgements. This underlies the reflexive researcher paradigmatic view and is a move away from views of linguistic competence. Chomsky views 'competence' as a psychological capacity and 'performance' as the realisation of an utterance. His view on performance differs as they centre around a grammar bound by an external process and performance relative to an idea of universal forms. Alternately, within these methods is an intention to foreground the reflexive qualities of discourse. This idea highlights agency involved in the production of dialogue and in turn highlights the cultural knowledge circulated through the medium of discourse (Johnstone and Marcello, 2010). Johnstone and Marcello (2010) note how Hymes created the SPEAKING model for researchers and those studying within the field of language variation research to study communication within fieldwork and language in practice. This work helped identify the components of linguistic interaction where one needs not only grammar and lexis, but also the knowledge to use words within context. In other words, communicative and linguistic competence. The purpose of this approach is the study of language in use and can be surmised in the following from Hymes:

> I should like to restrict the term "ethnography." I should like to give "ethnography" the connotation of inquiry that is open to questions and answers not foreseen, for which possible observations need not be pre-coded, and for which the test of validity need not fit within a pre-structured model. When anthropologists limit their inquiry to

> observations and questions for which the set of alternative answers is already fixed, I should like to say that that may be field work, but not ethnography.
>
> (Hymes, 1980:89)

The present study does not use ethnographic methods. However, Hymes' description (1964) of ethnographic approach is related to the present ethical reflexive approach as an attempt to position the study of language and variation as an emic study. Hymes sees the study of language as a rhetorical continuum or extension to the American linguistic anthropology of Edward Sapir who was seen as the founder of the science of linguistic anthropology. However, a linguistic division between approaches was formed as a functionalist perspective was introduced. This division relates to the SPEAKING model of Hymes' work which was in associations with Roman Jakobson (Hymes, 1975). Sociolinguists' concern with the contextualisation of language use was initiated and further developed alongside Goffman, Gumperz, and Ervin-Tripp who together established a research centre for urban study of language variation. Their pursuit of understanding is concerned with what a person knows about appropriate patterns of communication in his or her community and how they learn about them (Farah, 1997). This perspective offers an approach that looks at people's lives from their own perspective and from the experience of their own context (Hymes, 1975). In linguistic terms, these are approaches that promote an emic view of an individual's speech as part of a speech economy or 'communicative economy' (Hymes, 1964: 3) that speakers participate in as part of a social realm. Within this speech economy, speakers operate features that enter into their styles as well as the styles themselves. Therefore, what seems to be deviation or variation from a standard are structure and pattern from the point of 'communicative economy' of the group where these variations exists. Thus, linguistic variation can be studied from the individual's 'means of speech' (Hymes, 1964) with relative linguistic coherence from across the social context where these patterns of variation are found. Subsequently what the speaker can and does say can be contextualised within the communal and social realm. The word and experience of the participant is taken as the starting point for interactional communication. In this sense the qualitative researcher takes a bottom-up or emic view in order to gain perspective on macro-social or etic top down social views of social categorisation. This perspective enables a view of identity to establish within the interaction, whilst wider macro-social categorisations are viewed from within socially constructed linguistic processes.

A researcher that incorporates a reflexive approach that entails an emic perspective commonly takes consideration of ethical concerns and the position of the research in relation to the roles within the interview. This approach can be seen as part of a group of methods that emulate common skills of everyday life activities. The skills involved include asking and answering questions about people's concerns and thoughts on their everyday lives. Participant observations (O'Reilly, 2005) are one commonly cited data collection method and are concerned with the taking of notes during interviews and observations that are part of these interactions and was part of the research method used for data collection within this study.

Central to the interview process as part of an ethical reflexive approach is the endeavour to observe objectively. This, however, creates a tension between an attempt to capture objective observations whilst participating in an interaction as a fellow human being, a co-constructer of dialogue with empathy and sympathy to the perspective of the participant. The role of observation has been a contentious pursuit of sociolinguists in an attempt to record speech in its conceivably unmonitored or uninfluenced forms. Labov (1972a) argued for this concept of attempting to record forms that have not influenced by contextual or external circumstances when elicited by experimentation or artificial conditions. This it is argued, represents a paradox, as a style of speech where any systematic attention is given to its formulation will affect the language of most interest to the linguist. This is a central focus and concern for the present study. Labov's work developed the research approach for the sociolinguist through fieldwork techniques that enabled observation of vernacular speech, including interview questions (Labov, 1966; 1972b). Labov summarises the observer paradox by stating the sociolinguist's quest is to find out how people use language using systematic methods whilst not observing such behaviour in a systematically confined manner (Labov, 1972a: 209). Whilst this paradox will necessarily affect the characteristics of data collected, this paradoxical effect can be minimised. Through the adoption of a reflexive approach to interview, the observer takes part as participant creating arguably greater opportunity for the study of variation in practice.

A strategy for reducing the effects of the observer paradox is to gain the trust of participants (Labov, 1972). The relationship between the observer and the participant is a significant contributor to the information that is gathered during the interview process and an essential component of any fieldwork that entails a reflexive approach. In order for meaningful access to individuals' lives and experiences this trust needs to be gained and maintained throughout the interview and the lifecycle of the research process. Mistrust can be a response to approaches by figures of authority (Howley, 2016) and can also be a consideration as part of a reflexive

approach to data collection. A relationship built as part of a reflexive approach and process can be considered an essential aspect of the data collection and recruitment process and challenges outlined in 3.3. In order to be familiar with the participants and to bring ownership and agency of language and identity to the data collected, I had to partake as an invested participant and be familiar with community members that that they themselves associate with. I had to develop trust, and this could only be fostered through meaningful contacts established over a period of time and with trusted others who have a longstanding and deeper relationship than I could establish over my timeframe for this project.

The composition of the semi-structured interview is presented in the following sections. The structure of the interview process and the nature of each conversation is discussed in order to create a comparable degree of reliability. Before I write about the detail of recording of the semi-structured interviews, I will outline the value of a reflexive approach for researching language and culture focusing on ethnically diverse communities in a regional context. The analysis of lexical variation within interaction is central to the methodological approach utilised for this research project and allows for both analysis of the variation and its influencing factors. It also enables an exploration of identities that are products of the social interactions themselves, and how individuals are represented within these interactions, with lexical variation central to these interactions.

3.2.1 Accounting for Regional Language Variation

Many contemporary studies that focus on dialect variation and social and cultural factors that influence variation have a relatively small sample size (McColl Millar, Barras and Bonnici, 2014). This methodological practice is inherently characteristic of this type of language research. The richness of interview data means a smaller sample size allows for detailed and fine-grained analysis. This contrasts with earlier, traditional dialect studies of the 19th and 20th centuries which relied on large numbers of informants for quantification of results (Chambers and Trudgill, 1980; Labov, 1972; Lakoff, 1975). A limitation of these inferential studies is the underlying principle that a majority finding is inferred to all speakers, missing the richness and nuanced emic reality of the speakers within these communities. Studies that use interview data rely on involvement of the researcher in a social setting to uncover social and cultural factors that affect language variation from the standpoint of the community themselves (Wolfram, 2000). This type of linguistic research involves participant interviews and observation and is associated with relatively small sample sizes involving detailed observations of as few as four participants over longer periods of time (Eckert, 1989; Wolfram and Schilling-Estes, 2005;

Jones, 2012). Studies such as that by Jones (2012; 2019; 2020), through an interactionist approach, interpret language as meaningful within an interactive context, as a communicative tool and fundamental to the sociocultural linguistic approach of Bucholtz and Hall (2005). This approach is a guiding principle to the approach taken within the analysis within chapters 4 –7 and outlined in section 1.2.

Emic descriptions by participants in relation to interview data allows for the recording of complex ways in which individuals construct their reality. Language use can be viewed within the construction of identity through appearance, actions, and selection of discussion topics at the local level where meaning can be seen to emerge (Bucholtz, 1999). This approach allows the researcher to work with the individuals rather than as a researcher positioned as separate and aloof from the participants of the study. This is particularly important for this research, as research focusing on potentially vulnerable participants, or where the topic under discussion is particularly sensitive, participant observation over extended periods can reduce the sense that people are being put under the researcher's microscope, thereby reducing the perceived power differential between participants and the researcher. This approach aligns with other sociolinguistic studies focusing on individuals as interactionally involved, such as the seminal sociolinguistic research of Eckert (1997). Eckert's work in *Vowels and Nail Polish: The Emergence of Linguistic Style in the Preadolescent Heterosexual Marketplace* (2006) applies and develops Lave and Wenger's (1991) theory on communities of practice where speakers share pursuits and similar positions within a social context. The theoretical perspectives on communities of practice also involve the standpoints of performativity theory (Butler, 1988;1993) outlined in section 1.2. Whist not technically applicable to this study, the model is of relevant note. Previous sociolinguistic research has focused on sociological categories that influence variation such as education, socioeconomic class, income, values and types of residence such as mobile or static. Labels such as 'Gypsy', 'Roma', 'Traveller' are explored within the emic descriptions given by those who identity within these broad social categorisations. To this end, the participants create their own descriptions, positioning their linguistic practices as part of this identity. These polarised perceptions have implications for the social identity and integration of community membership.

Within the study, the language that is identified by participants as relative to their self-ascribed identity and ethnicity is referred to as a variant. These linguistic features are referred to as a 'mixed-language' and are also labelled as Anglo-Romani (Matras, 2010). That which has been referred to as Irish Pavee Cant is also referenced throughout the analysis of this thesis as 'non-standard'. Howley (2016) describes how Roma are frequently treated as an underclass or

outsiders of the social class system. It is therefore within this context that the research questions are explored and answered within my own reflexive framework for enquiry.

3.2.2 My Position as an Ethical Reflexive Researcher

As with Howley (2016) and Eckert (1989) I will position myself via self-description as part of the framework for analysis. In research positioned around the principles of social and ethical reflexive research, it is often cited as necessary to acknowledge the background of those carrying out the investigation. It is important to recognise that the research is partly a production of ourselves and what we bring to this research, whether it is embedded empirically in quantitative analysis, qualitative endeavours, or an ethical reflexive approach to research (Hymes 1975). There is therefore a significance to acknowledging that personal and individual characteristics influence how we study and consequently what is found. The tenets of this method influence the approach that an ethical reflexive researcher can hope to maintain within social interactions.

> …there exists no neutral position for a researcher - if you are engaged in social interaction, you are part of that interaction, and who you are is going to affect the kind of data you have access to.
>
> (Modan, 2007:286)

In view of this, those researching language variation need to acknowledge the division between participation and observation. In relation to this tension Tedlock (1991) highlights the importance of reflexivity within a methodology. The approach outlined is one where the researcher self-explores the participation itself and of themselves as researcher in order to understand the context of participant observation by what can be described as self-observation of the participation. This creates a richer understanding of the perspectives developed out of the research. The principle behind this endeavour is to gain an awareness of the biases, experiences and interests that will influence the way events are experienced and how they are observed. This then forms a richer tapestry or depth of information with which to form the analysis. As is common practice for a study of this nature therefore, I position myself as a researcher within the context of my own unique situation to broaden the framework of this research methodology.

I was born in Worksop and was raised in Nottinghamshire. For the first five years of my life my family lived in a rural location. First in a small village, and then moving to a dilapidated

farm in need of complete renovation where we lived in a static caravan over two snowy winters. I have vivid memories of rural life during this period and recall playing in the fields, taking outdoor baths in a wooden barrel filled from a well in the yard on warmer days, and exploring the surrounding countryside. This was always a favourite time which entailed eating fresh vinegar leaves, picking mushrooms, and gathering freshwater cress growing wildly in the streams. I have vague memories of the caravans and older style 'Gypsy' wagons pulled up on nearby small laybys and sheltered areas near or along the roadside. This had always seemed very curious to me as to who these people were and what these wagons were doing parked along the roadside. I was raised in a market town and left for the city and travelled and lived overseas for some years. I returned to the town where I grew up and where I again became motivated to understand more about the culture and diversity that influenced myself and those I grew up alongside. As with related approaches that consider language variation in context (Howley, 2016, Eckert 1997) these details outline my own background and notions that influence the decisions and interpretations that are carried within the analysis of data collection using the methods detailed here.

3.3 Sampling

The project used purposeful sampling (Creswell and Plano Clark, 2018). This involved discussions with relevant professionals, friends and family to recruit individuals that self-ascribe within the identity appellations outlined in Figure 4.3 broadly pertaining to a GRT identity. The rationale for this sampling was to identify those participants who identify or self-ascribe with a GRT identity. I contacted Liaison Officers from local councils and these contacts introduced me to individuals who self-identified as Roma and Irish Traveller.

Opportunistic and emergent sampling (Patton, 2002) were also used as decisions were made as situations unfolded spontaneously. This was used in situations where a friend was introduced through a spontaneous conversation which led to a longer interview. Justification for this inclusion is based on self-ascribing within a broader GRT identity. In this instance, I was introduced to someone during a visit to a local business together with a family friend. The family friend was an acquittance of the business owner. The business owner suggested we talk with his friend who identifies as Roma.

An opportunistic sampling method allows for reflection and considerations that can be made based on preliminary findings. This method of sampling was adopted due to the challenges that are an inherent in identifying and selecting individuals as part of a minority group. Again, the justification for this sampling is based on the individual self-ascribing within the broader GRT

identities. Chain sampling or snowball sampling was also component of the sampling method (Holmes and Hazen, 2014) whereby participant numbers increase over time through network connections.

3.3.1 The Participants

There were a total of six participants (all names used are pseudonyms) who took part within the research project. These were selected based on their self-ascribing to an identity within the broader identity label of GRT. A case sample of six was considered beneficial in terms of the amount of interview data that was collected. This allowed for a finer more detailed analysis of interview data and analysis of background factors represented within the six participants.

Whilst recruiting participants, I mitigated for sampling bias by requesting male and female professionals to actively recruit to encourage both male and female and a mix of age ranges to participant in the research. The limitations in terms of gender representation and age variation is outlined in section 3.8.2. Tabulated below in Table 3.0 are participants details with information on the context within which the interview took place.

Table 3.0 Participant information

*Participant	Context	Data collected	Length of interview	Demographic information	Region and Ethnic background
Judy:	Judy was introduced to me by a professional Liaison worker who is employed to improve relations between the local community and provide support for members of the community where necessary. There were three people present at the interview. Interviewer, Participant and Liaison officer. The Liaison officer did not take part.	The interview was not recorded. Field notes were used to aid reconstruction of the interview and lexis was transcribed as with others. Semi-structured interview Identity Score index Social Network measure SRN questionnaire	The interview duration was approximately 30 minutes and support workers also took part at intervals.	Judy was approximately 35 to 45 years of age and demographic data was not volunteered.	The interview took place in Leicestershire and Judy explained that she is from this area. Judy described herself as having an Irish Traveller background.
Graham:	Took part with a close family friend at the friend's home. The friend also participated in the interview.	The interview with Graham was recorded and transcribed. Semi-structured interview SRN questionnaire	The interview with Graham took place over a duration of 45 minutes.	Graham is in his 30's and a tradesman.	The interview took place in Derbyshire. Graham explained his identity as being part of the 'Gypsy' community, having knowledge of language and culture and integrated into this community's way of life
Duncan:	Duncan lives with his family and was introduced to me by an acquaintance. Duncan explained his	Duncan was particularly helpful and responsive to correspondence both	The first interview was 45 minutes and the second one and half	Duncan is 40 – 50 years of age with older children and a wide	Duncan lives in Nottinghamshire He is a local businessman and trader. Duncan was raised in

	motivation for sharing his culture and ideas about his community and shared photos, anecdotes and ideas about language to help with this research. He invited me into his home and shared cultural information that was relevant to the research. Duncan and the researcher took part in the interview only.	before and after interviews. The interviews were both recorded and transcribed. Semi-structured interview Identity Score index Relation analogue scale Social Network measure SRN questionnaire	hours. The interviews took place over 2 separate occasions at his home.	network of friends, family and acquaintances.	Yorkshire and discussed travelling across the UK as a child with a father who spoke Romani and a mother who spoke Irish Cant.
Clinton:	The interview was set-up spontaneously through an acquaintance of his after answering a telephone call related to the possibility of a discussion about my research and was able to meet for an interview The interview took place at his home and took place initially with his wife and sister-in-law at the start of the interview. Another acquaintance that helped arrange the interview stayed for the duration of the interview whist the sister-in-left left shortly after the start	The interview was recorded and transcribed. Semi-structured interview Identity Score index Relation analogue scale Social Network measure SRN questionnaire	The interview took place over 2 hours.	Clinton is 50-60 years of ag. He has lived in the area since childhood and was raised speaking Romani by his father.	Clinton resides in Nottinghamshire.

Jake:	He was introduced to me by an acquaintance. Jake invited me to take part in a conference he had organised that was to help promote the awareness of the Gypsy, Roman and Traveller culture.	The interview with Jake took place by mail. Semi-structured interview Relation analogue scale Social Network measure SRN questionnaire	The interview took place over several letters each one with different sections of the research data tools.	Jake is 60-70 years of age.	Jake was raised in the Lincolnshire. He is an active member of the community in that he promotes cultural awareness as well as the language as part of this cultural awareness raising. He considers himself part of a wider network of Romani speakers
Ben:.	Ben was introduced through an acquaintance and was motivated to share his culture and heritage to help raise awareness as an active member of the community. Ben's interviews took place through email.	Semi-structured interview SRN questionnaire	Over several correspondences.	Ben is 55-65 years of age.	an active member of the Romani community that promotes culture and language. He was raised by a Romani speaking father and this he makes clear had an influence on his current use of his own Romani He lives now in New Zealand and lived in the UK as a child and considered his Romani. heritage attached to Wales and the Welsh Romani.

***All names are pseudonyms to retain anonymity.**

As noted in Table 3.0, there were other notable participants who took part within the interactions. These interactions are explored within the analysis. As mentioned in section 3.3 as part of the collection method, spontaneous sampling was used. Table 3 indicates the group dynamics for each interview and the interlocutors that took part. Whilst this added variation in terms of the collection method, this further enabled authenticity and facilitated the conversations where had this been less fluid, discussion around lexical variation may have seemed forced or inauthentic.

Participant Ben does not live in the East Midlands and therefore do not represent an individual case study for lexical variation within the region. However, for the purpose of comparison it was considered an effective means to contrast findings based on his lexical variation and analysis and discussion of background factors and measures alongside meaning in use. It was assumed this would add detail an depth to the analysis that would otherwise have been missed.

3.4 Collecting Language Data

This research design has been adapted from dialect research methods that have been carried out in the UK and with a specific focus on the fieldwork done on lexical variation within the regional communities of the UK (Llamas, 1999). As outlined, the approach to data collection is founded on the tools developed as part of the Survey of Regional English (SuRE) methodology. The semi-structured interview is also conducted and analysed in the tradition of sociolinguistic research where linguistic resources gain social meaning through interaction (Bucholtz and Hall, 2005) outlined above and in section 1.2 and positioning this work within the framework of sociocultural theory of identity and linguistics variation. The details of the research instruments are discussed in the following section on methodology. A mixed methods design has been adopted for the purposes of this research project.

From the literature, evidence suggests that social networks (Donghui, 2012; Milroy, 1982, 1987; Howley, 2015; Velazquez, 2013) and attitude (Zhang, 2009, Suek, 2014, Gharibi and Boers, 2017, Weyers, 2014) contribute to the maintenance and shifts of minority variants and linguistic forms that co-exist within a standard form or language. The research methodology is therefore designed to identify dialect variation, in addition to an analysis of variation positioned within social context. Responses were elicited via semi-structured interview which entailed several components outlined in the sections 3.5 – 3.7. Conversations were transcribed and analysed for thematic content using NVivo qualitative data analysis software. The use of NVivo allows for mapping and identification of thematic content. This enabled a structured

and systematic method for identifying and analysing dialogue. Selected excerpts are provided within the data analysis and discussion chapters 4 to 7.

3.5 Lexical Methods

Comparability of lexical variation has been made both regionally and between groups and individuals using Sense Relation Networks (SRN) as the primary data collection instrument. This collection method was developed by Llamas (1999) as part of a sociolinguistic study of Teeside English and part of a wider project, the Survey of Regional English (SuRE) (Kerswill, Llamas and Upton, 1999). Four SRNs were used as part of the interviews. The SRNs were developed from a pilot investigation. The pilot study is outlined in section 3.5.1 below.

3.5.1 Piloting of SRNs

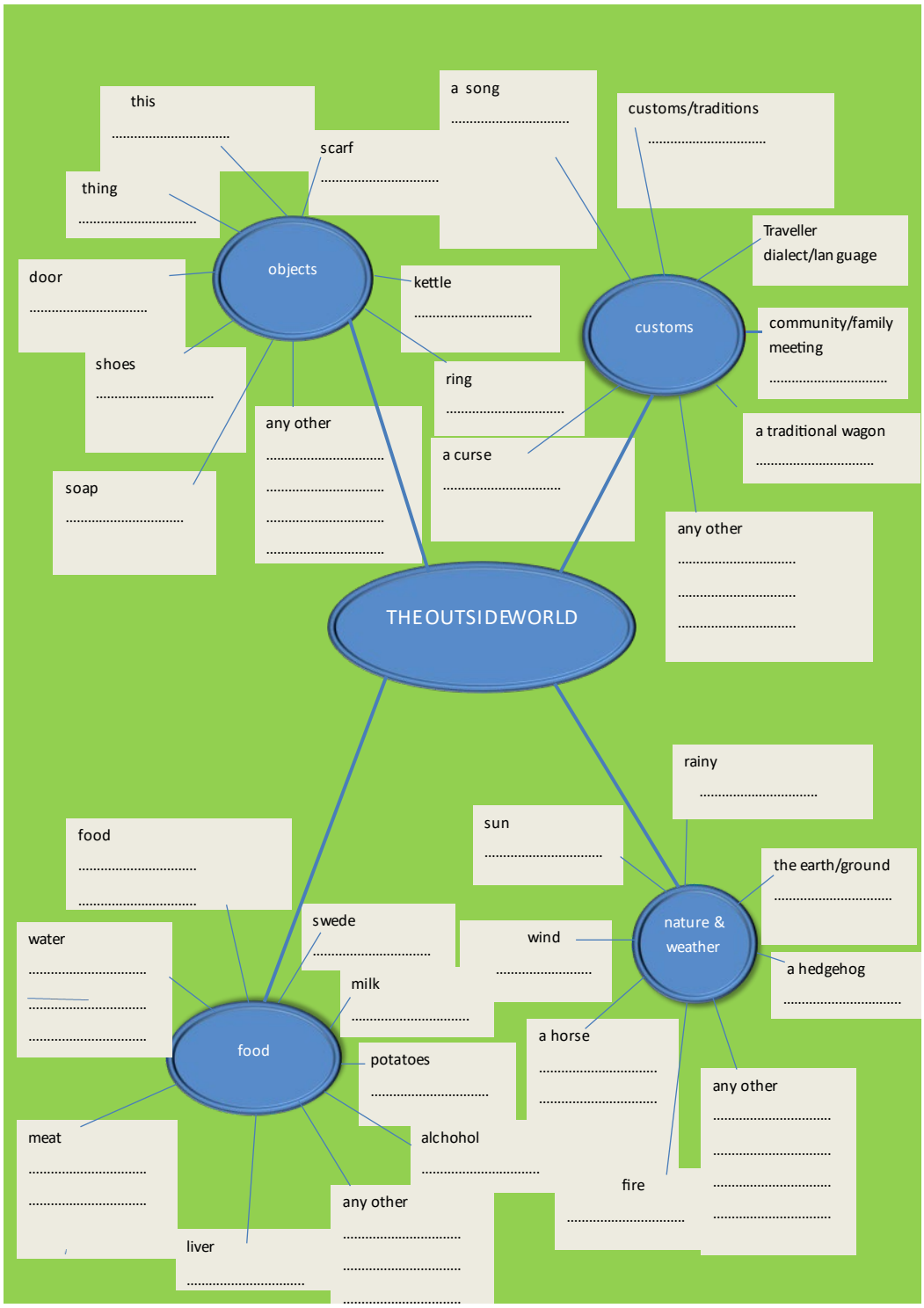
These words chosen for the SRNs were initially piloted. The selection of semantic fields and standard notion words in the four SRNs shown in Figure 3.1 and Appendix 1 are the result of this initial trial. Trialling consisted of a short interview asking a participant who identified as having a Roma background, whether they were aware of variants. The pilot study consisted of a list of words that were thought to potentially elicit vocabulary that relate to Anglo-Romani and Irish Pavee Cant lexis. The list of English words was developed from a list of Anglo-Romani and Irish Pavee Cant words collected from three sources (Dawson, 2002; 2009; Le Bas, 2018; Manchester Romani Project, 2021). This approach was both practical in terms of time and served as an identifiable starting point for the purpose of piloting the instrument. Both word lists were translated into English and categorised into semantic domains. This prototype SRN was then used with a participant to test its ability to effectively elicit words from a participant who self-ascribes to a GRT related heritage or identity but who was not part of the main study. Those words that did not elicit variants were dropped or replaced. The primary target was to create conditions for maximum elicitation and therefore a richer collection of lexical variation.

3.5.2 Sense Relation Network

The principal design of the SRNs allows for visual linking between linguistic forms and assists in accessing lexical items due to the hierarchically organised structure of the mental lexicon. The SRN is a visual representation of a semantically connected group of words. A central or node word is placed in focal position and words that are connected in meaning are linked visually with lines that branch out from this central point. The interviewee is asked to produce associated ‘non-standard’ words that are represented with ‘standard’ forms within this hierarchical depiction. The rationale for the use of SRNs is in consideration of how words are

theoretically stored and retrieved, being relationally interconnected within the mental lexicon and metaphorically referred to as a web of words by Aitchison (1997). The use of the SRN allows for flexibility during interviews, and words and phrases can be added or left out as and when appropriate. Figure 3.1 shows the SRN 'The Outside World' used in the collection of lexical variation for this study.

Figure 3.1 Sense Relation Network: The Outside World



The notion words (central to the SRNs) did not change whilst lexical items ('standard' English words) have been edited. 'Standard' notion words with no responses have been deselected where no variant has been found. Each SRN provided room for variants of items, which were not listed in the SRN that, therefore, allowed for insertion of words from outside those items on the SRN. The SRNs showed potential to represent geographical and individual variations of those that identified within a GRT identity or heritage from the East Midlands to gain a clearer conception of what variations there were within the languages that are associated with GRT identity and heritage.

The organisation of lexis around interconnected semantic fields gave structure to the interview process and allowed for a conversational approach, which would not have been possible had the words been presented in an unstructured or formatted list. Each semantic field, therefore, presents a potential topic for discussion related to dialect use. Each standard notion word has space below it for the participant and researcher to insert a non-standard dialectal alternative. Each SRN is printed in a different colour to increase visual impact in order to engage with participants by encouraging them to complete the task.

The use of colour, including visual aspect of the SRN notably improved the engagement of the participants. Whilst without comparative evidence, intuitive accounts and relating to feedback from interviewees, it was shown that with use of the SRN, discussions had a firm tendency to focus on variation and to focus around the given semantic fields. This was beneficial, in terms of comparability across the interviews, and for the facilitation of elicitation. The level of spontaneity was enabled through the associations of notion words which increased time efficiency, and avoided a formal interview style of questioning (Llamas, 1999). This has also been of significant benefit in increasing authenticity of responses. As this design corresponded with previous dialect surveys (Llamas,1999) from the groupings of questions in the Survey of English Dialects (Orton and Dieth, 1971) there was a degree of comparability with the findings of this research project. This allowed for a measure of reliability to be drawn from the results of the research findings of this study. The semantic fields were based around the categorisation of lexis, which had been established as a means of eliciting variation. This was standardised through the piloting of an initial set of words. There were four fields, that were: 'The Outside World 1', The Outside World 2', 'People', and 'Feelings, Actions and States. For a full list of the lexis and the layout of each see Appendix 1. These fields were developed through a trial to relate those lexical items and collate them into four semantic fields

as outlined in section 3.5.1. The result was found to be the most logical in terms of semantic categorisation.

A version of the instrument was previously used as part of the Teesside English project where the administration technique meant the SRN was received five days prior to the interview as a means of avoiding unease, and to promote higher yields of response (Llamas, 1999). The methodology deviated from the SuRE method on this point. A decision was made to interview without this step in the case if four particpants, as spontaneity was considered more valuable in terms of gaining richer results. The circumstances for Judy, Graham, Duncan, and Clinton's interviews did not allow for prior communication. However, Jake and Ben were unable to physically meet face to face. Therefore, the decision was made for them to have the SRNs as part of the correspondence owing to practical circumstances. This was discussed as a limitation to the study (section 3.8.2) and a consideration in terms of comparability between Jake and Ben's variation within the analysis of chapters 3 through to 7. Judy, Graham, Duncan and Clinton responded without prior correspondence, with their lexical variations being considered in respect of this methodological constraint.

Participant literacy levels were also a consideration and was raised in the cases of Judy and Duncan. This meant participants have had their responses transcribed by the researcher at the time of interview. This was again a consideration as part of the analysis and comparison of lexical variation observed within chapters 3 through to 7.

Consistency was managed for data collection for each participant. The administration was consistently maintained between cases and benefitted from a coherent focus on the data collection instruments. I elicited and wrote down the notion words onto the SRN for the informant in the case where the SRN had not been completed prior to the interview for Judy, Graham, Duncan and Clinton. Interviews were shown to yield dialectal variation (Llamas, 1999) with particular focus on lexis and this has proved the SRN to be an efficacious instrument for the purposes of this study.

3.5.3 Archive Sources

The archive sources are listed in table 4.5 in the first instance (see appendix 9 for full bibliographic references). These sources represented a broad range of Roma and Irish Cant literary sources in which lexical variation was documented. Each source was coded with letters, or with both letter and number (i.e. MC). If a lexical variant used by a participant corresponded with that found in an archive source, the corresponding code was indicated below the elicited

lexical variant. In terms of lexical representation and historical documentation of variation, a novel method was necessary for addressing the disparity between sources of reference. The sources used for historic comparison used a variety of transcription and notation methods to document lexical variations. This and also omission of transcription methods was also notable within historic records. Also of importance was the variation found between participants in terms of their orthographic representations of lexis. To address those underlying factors relating to spelling differences, a method for comparing phonological comparison relied upon judgments. Those judgements relied on the association of consonant and vowel patterns that were considered relevant and relatively similar. Distinction between consonants were less ambiguous. However, vowel transcriptions and their inherent ambiguity means researcher judgements were necessary. These researcher judgements were based on contextual factors, such as, the linguistic environment in order to distinguish phonological features that were of significance in relation to variation uncovered within the analysis.

3.6 Interview Methods

The interviews were guided by a list of questions that formed the Identity Questionnaire (Llamas, 1999). The research tools including the Identity Questionnaire are given in Appendix 2 - 4. The Identity Questionnaire defined the interview as semi-structured and guided the conversation. Similar to the administration of the SRN, the Identity Questionnaire had in previous studies (e.g., Llamas, 1999) been given to the informant five days prior to the interview. This was a limitation in terms of the present study and effects for inconsistencies were mitigated and acknowledged as part of the process of analysis through chapters 3 to 7. The researcher and participant's discussions were open, and explored lexical variation as part of the conversation. As this was the case for four of the participants there was consistency within this approach. This was advantageous as a means for eliciting authentic data, although time efficiency allocated for the interview was an additional factor. In the case of Jake and Ben's cases, their data were considered in respect of this methodological limitation within the analysis.

The questions designed for the Identification Questionnaire were those that had been used by Llamas (1999). These were adapted and developed to reflect the context and purpose of the present study. The questions as a component of the semi-structured interview were used as guiding topics rather than rigid boundaries for discussion. Conversation was therefore a much more unrehearsed event with much richer data as a consequences for the analysis. The ideas of the questionnaire as a principal basis for discussion in instances where SRN responses elicited

limited responses was supported and showed itself to be an invaluable part of the inventory. The design of the semi-structured interview was made to gather information to help explore notions of identity and the participants' self-perceptions and their place within a social network by engaging with community members themselves (Cohen, 1985). The design of the questionnaire involved questions centred around but not limited to the topics of identity (Fishman and Garcia, 2014), age (Chambers and Schilling, 2013; Pichler, Wagner and Hesson, 2018) gender (Eckert & McConnell-Ginet, 1992) and concepts concerning social networks (Milroy, 1987) to understand attitudes and ideas on those areas to contextualise language and variation data with a focus on those cultural perspectives. The notion of social network applied here to refer to an individual's perception of a wider group of speakers they identified culturally with. Social network was also measured by self-perception as participants were asked to report their ideas concerning social network density (see limitation section 3.8.2). Example of questions from the Identification Questionnaire (see Appendix 2) are:

> *Is there a difference between the way older and younger people speak within the Traveller or Roma community? Can you describe some of these differences?*
>
> *Do you change the change the way you talk depending on the situation? If so, in what situation, and why?*
>
> *What image or description of the Traveller or Roma community would you give to someone who didn't know about it?*

The discussions focused on language, and ideas pertaining to community and attitudes within the region as well as social labels such as GRT to enable exploration of such notions.

The Identity Questionnaire acted as a 'safety net' when conducting the interview as it gave purpose and guidance to the conversation when and if necessary. Experience from conducting the interviews showed that it was necessary to engage a participant in unconstrained discussion rather than slavishly follow t questions on the Identification Questionnaire itself. The success of the interview relied, therefore, on the researcher's ability to guide the conversation towards topics and questions from the Identity Questionnaire for the sake of comparability between other participants, although without limiting the discussion to those of the questionnaire alone to better facilitate an authentic generation of participant responses.

3. 7 Quantitative Methods

A mixed methods design was chosen to enable the comparability of a quantitative measurement of attitudinal ratings as reported by the participants in the study with a structured qualitative analysis of conversation between the researcher as well as other participants, with evidence of lexical variations that were self-reported by the participants taking part in the research project (Creswell and Plano Clark, 2018). The purpose of the numerical measures ensured the analytic focus was on the data. Regionality, ethnic and heritage labels were identified by the participants. This encouraged discussion about lexical variation to be contextualised in relation to ethnic heritage as well as other social parameters. The initial principle behind the use of quantitative data was, in part, to inform and counter the effect of a top down approach to social categorisation. The use of an Identity Score (Appendix 3) index was to acquire an understanding of social category and identity from the speaker's perspective and as a quantifiable means of measuring an individual's attitude towards language and identity. An example of a question from the Identity Score Index is:

1. *If you were on holiday and saw someone you had never seen before but thought they came from the Traveller or Roma community (e.g., you overheard their accent and recognised it, if they were discussing Traveller or Roma community related topics etc.) would you:*

    a. *Feel compelled to go and ask where they were from and strike up a relationship*

    b. *Feel you had something in common but not do anything about it*

    c. *Not feel any differently than you would towards any other stranger*

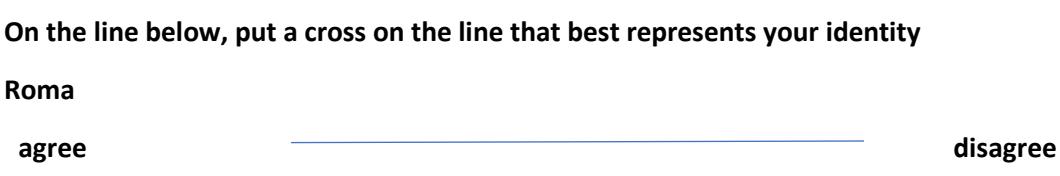
The index had been developed by Llamas (1999; adapted from Underwood, 1988) and further modified for the purposes of this study with specific regards to the original devised by Underwood (1988). Underwood developed a score index as a means of assessing regional Texan affiliations by comparing the use of particular local variants. The ideas for the index were based on Le Page and Tabouret-Keller's (1985) theory on acts of identity. This study has used an extended adaptation of the Identity Score Index. The index was originally conceived to examine notions of ethnic identity, and, therefore, the present study's extension of use was such that questions were extended to include identity scores for local or regional identity, including also for ethnicity. Two numbers were recorded for each participant (see Table 4.2). One identified the regional index score, and the other for the ethnic index score. Graham, Jake,

and Ben did not complete this component of the SuRE method as they did not return this section of the survey. This again is discussed as a limitation of the present research as absence of this data didnot allow a full picture of comparison. This was a consideration as part of the analysis within chapter 4 through to7. The questions for the Identity Score Index were generated about regionality and ethnicity and were designed to enquire how informants might respond towards those from their local area as well as those from their own self-ascribed ethnic identity.

There were a total of 45 questions in the semi-structured interview, whilst only a limited number would consistently be used for each interview. The Identity Score Index had 10 items and was therefore quite short. The short administration for the questionnaire meant that it was practical and efficient to use. The questions involved scenarios that were explained verbally involving hypothetical scenarios. The scoring of the questions was quick, and answers were recorded as scores 1, 2 or 3 with 3 representing a hypothetical scenario suggesting strongest feelings of allegiance to ethnicity or region, and was easily noted by the researcher. A score of 3 would be given for the answer *a* (above) as it showed the closest alliance to a notion of cultural relation. As a quantifiable measurement of attitude, this was an effective method to provide a quantitative measure as part of the analysis, by integrating interview and SRN data.

Visual analogue scales (VAS) (Appendix 4) are visual depictions of relationships to notions of regional and national identity. This study made use of this quantifiable data collected through the use of VAS (Llamas and Watt, 2014). These VAS have been adapted for data collection administrable via laptop. Paper administration was used where a laptop was not practical. The scales allowed for measurement of relationships of notions concerning self-identity. The principle behind the use of the scale was to avoid problematic tendencies associated with Likert scale composition, such as, a central tendency bias (Sims, 2002). The use of this scale denied a person the opportunity to select a central number, such as, in the case of a scale from 1 to 5, where 3 was commonly selected. Bespoke software (Braun, 2019) converted the position to a numerical value for quantification of the measurement for efficiency shown below in Figure 3.2:

Figure 3.2 A question from the relational Analogue Scale



Here, the participant was asked to draw an X on the line. The question and drawing of the line was done via a PC, by automatically calculating a representative number. The visual nature and ease of administration engaged the informants and facilitated the speed and ease of data collection at the time of the interviews. For the purpose of the study, direct comparison of informants' identities and how they might have promoted one over another was not considered as quantifiable and measurable, and the measures indicated by these methods were indictive for individual factors of influence upon variation. Whilst not a representative numeric, the measure was used as a means for analysing individual influences of variation in terms of ratings of individual, regional, and national identity. This was more substantially an overt indication of participants' self-ascribed identity. It should be noted that the software (Braun, 2019) did calculate a numerical score.

A rudimentary measure of self-perceptions of social network density was also included within the study. The questions asked the participant to estimate the percentage of neighbours and friends within the area that they considered as part of their GRT heritage group. This was again used as a means for analysis influencing factors of variation between participants. Limitations of self-perceptions are acknowledged as part of section 3.8.2 as self-reported data can be subject to biases such as idealisation.

3.8 Theoretical Framework for analysis

The study's focus, in addition to exploring lexical variation, is the meaning of lexical variation within a social context. The framework of Bucholtz and Hall (2005) was selected as a basis for exploration of ideas of identity as social practice within interaction. They outlined five principles (below) that were involved in the construction of identity through interaction. They outlined how identity was the product rather than the source of linguistic practice. They suggested social categories were subject to taking particular stances and were emergent within interaction. In this sense, identities could be indexed through labels, through stance taking, use of styles and linguistic choices, such as the use of pronouns. They argued that meaning was relationally constructed, so, for example, whether difference or similarity was highlighted within a discussion. They also proposed that meaning was a social act and part a negotiation. A theoretical understanding of social meaning and linguistic variation was associated with the idea of indexicality (Silverstein, 1976, 2003; Ochs, 1992) where social interpretation and linguistic forms interacted.

The analysis for this study explored the social meaning of lexical variation within interaction through observations of 'stance' taking (Du Bois, 2007: 163). These observations were based on interactional dynamics and centred around discussions of lexical use. These qualitative analyses were conducted against the backdrop of the quantitative background data. The study explored lexical variation and the social meaning of lexical variation within interaction. This provided a fuller picture of lexical variation within its social context. The analysis of stance within the chapters 4-7 are featured throughout the discussion of lexical variation. This additional layer further explored factors that influenced variation. Rather than an analysis of stance per se (Keisling, 2009; Bucholtz, 2012), its function was to indicate participants' attitudes and evaluations towards their identity and their use of lexical variants. The analysis explored meaning that was identified through both inference and speaker utilisation (Hall-Lew, Moore and Podesva, 2021). The social meaning of lexical variation therefore, was the focus for this aspect of the analysis.

3.8.1 Ethical considerations

Ethical approval was applied for and given through Nottingham Trent University ethics committee and were an important aspect of the study. The ethics committee's recommendations were adopted and these informed methodological decisions during the process of data collection. Data collection and data management followed university ethics policy, data management policy, including regulations for the use of storing sensitive data digitally, and those guidelines were followed for risk assessment in the research. As part of the data collection a complete data management plan (see Appendix 7) was developed and approved by the university data management team.

As part of the ethical approval procedure, a Participant Information Sheet was developed. This outlined details about the project and was given and explained to each participant before they took part. The Participant Information Sheet can be found in Appendix 5.The information sheet focused on several main points. It was used to give individuals a detailed understanding of the purpose of the study, why they had been invited to take part, what would happen during their participation, and how the information would be used. The information sheet also specified that they were free to withdraw from participation at any stage of the research.

In addition to the information sheet, each participant was asked to consent to their participation in the study (See Appendix 6 for the consent form). The consent form consisted of seven areas and related to that given in the information sheet. The consent form indicated that the

participant would be recorded, and the data would be used for the purpose of the study. Participants understood all information would be anonymised and they would be free to withdraw at any point.

3.8.2 Limitations

Variation studies using the SuRE method (Llamas, 1999) used larger sample sizes than the current study. This study used a case study of six participants. The reason for using a smaller sample size related to the study's focus. For the purposes of this study, the identification and analysis of variation and meaning in context was the focus rather than macro-social patterns of variation. The findings related to individuals' variation rather than macro-social patterns of variation. Related studies with smaller sample sizes have shown they could make detailed commentary on social meaning and linguistic variation (MaColl Millar, Barras and Bonica 2014) providing support for the use of smaller case samples.

The study referred to a number of archive sources (see appendix 9). These varied in terms of how the data was collected, such as, the use of identity measures or transcription methods. For example, the archive sources included an online archive (Manchester Corpus, 2021) of Anglo-Romani word lists. These included reference to 'folk' word lists, such as, Dawson (2002). These texts were not peer reviewed but were produced by individuals who were enthusiasts or have been described as 'Gypsologists'. The collection methods are outlined within these publications although attention to phonological transcription varied between sources and between transcription methods. Within this study these archive sources were used for comparative purposes. However, the academic rigour in their production is a limitation acknowledged as part of the study and a consideration within the analysis and discussion that followed.

The interview circumstances also differed between participants. This varied in relation as to whom took part in the interview (outlined in section 3.3.1) and the circumstances, such as, the times and length of each interview and not being able to send out information sheets prior to interview as proposed within the methodology (Llamas, 1999). Consequently not all participants were able to complete all aspects of the SuRE method. This was evident in the omission of data from the background information presented in sections 4.1.1. This was a consideration within the analysis and discussion and reflected a limitation in terms of the comparability between individuals' variations. This, together with the comparability between the results of this study and other studies that used the SuRE method indicated that comparison was possible.

Sampling bias was mitigated by asking others to assist with recruitment. As many of the participants that came forward were older males this approach was not sufficient to mitigate for sampling bias completely. It is acknowledged that the majority of the participants were from a similar age group and only one female took part in the study. This is acknowledged and a consideration as part of the analysis and discussion of the study. It was also significant that the findings were not, therefore, intended to reflect a wider sense of the GRT community.
The notions of identity, social network density, and regional and ethnic identity scores were based on individual perceptions. Individual perceptions could be based on self-ideals and not necessary realities that could be objectively observed. Therefore, whilst this was central to the focus of the study, it is acknowledged that discussion and conclusions were not inferential and did not represent wider macro-social characteristics of language variation and meaning. It should be noted that the notion of Social Network (Milroy, 1987) as an influence on language change and maintenance was used broadly and exploration of the concept was based on self-reported measures.

3.9 Summary

This methodological section has outlined the research methods used for data collection. The section started with an introduction and rationale for the ethical reflexive research methods that have been used. This section outlined the purpose behind the use of this method as a tool for qualitative research. This was established by reference to the literature and contextualisation within the research of minority language and variation. This chapter then outlined my own position, view, and background to determine and highlight my own ideological position as integral for the analysis of this investigation. Here I discussed my own experience, moving on to a description of the six participants, who took part in this research. The methods for conducting this research were then described, focusing on the elicitation methods for lexical variation which was central to the exploration of language and variation within this study, with the interview method central to the focus of an ethical reflexive approach for data collection. Quantitative measures were described that enabled participants to self-report their attitudes towards identity. This constituted identity measures adapted from the SuRE method and a means for observing individual factors that were influential on variation. Finally, ethical considerations and procedures outlined were followed by limitations that applied to the analysis chapters 4 through to 7 that preceded this methodology section.

Chapter 4 Analysis: Outside World 1

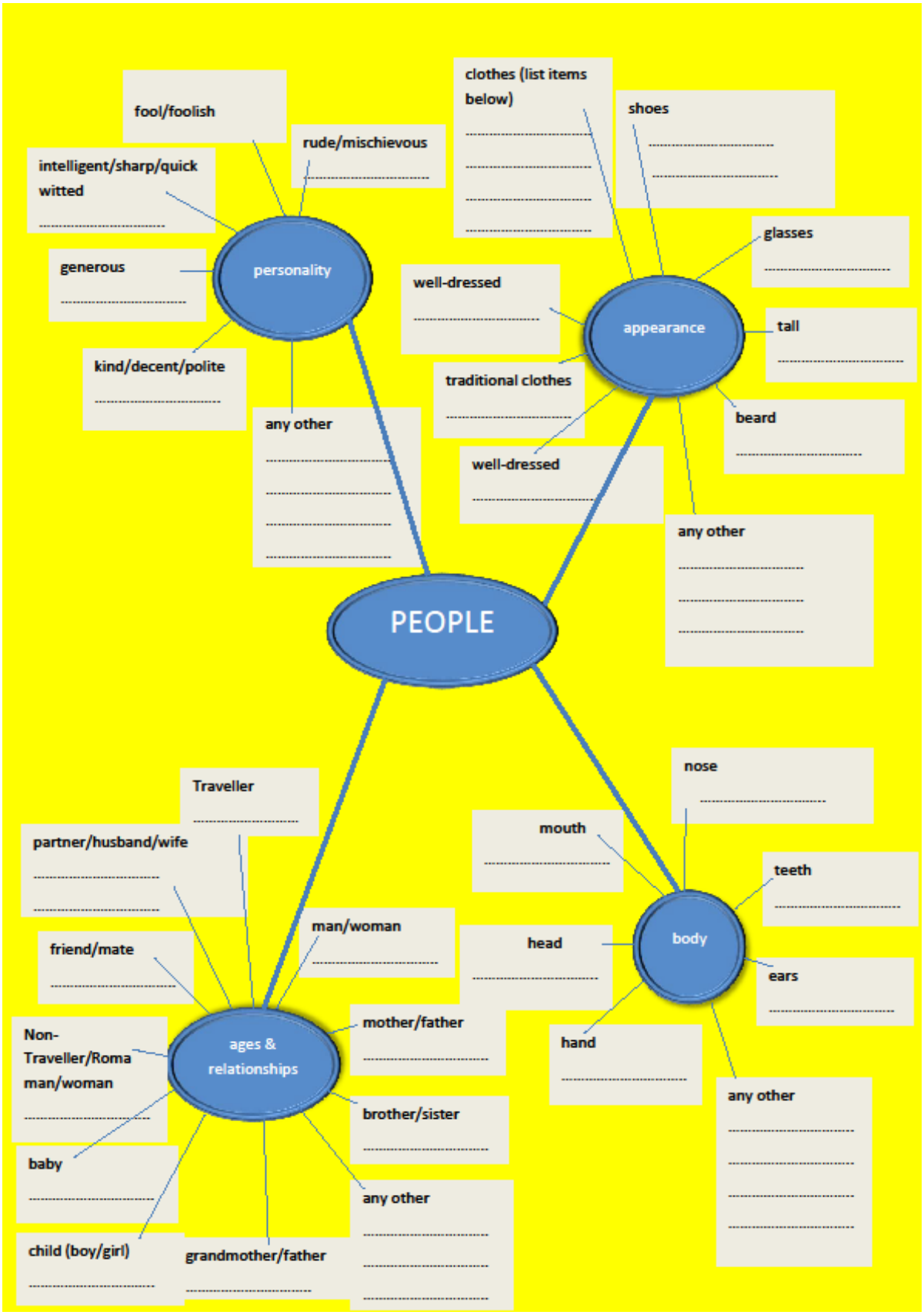
4.1 Introduction to the Analysis

The following chapters 4 – 7 are a detailed analysis of the lexical data together with identity data elicited by using the SuRE procedures for collecting lexical variation. The SuRE method used Sense Relation Networks (SRNs) to collect non-standard lexical items from participants. These were used with semi-structured interviews and questionnaires that measured attitude and identity discussed in the methodology chapter section 3.7. See section 3.5 for further elaboration of the procedures for development of the SRNs. There were in total four separate SRNs used for this study. Each SRN represented a superordinate semantic category. These are:

- Outside World 1
- Outside World 2
- Feelings, Actions and States
- People

These can be found in Appendix 1. An example is shown in Figure 4.1 which illustrated the SRN for 'People' showing semantically related categories: 'personality', 'appearance', 'relationships' and 'body'. The example revealed how the semantically related categories were visually connected. The method for developing each category (3.5.1) within each SNS is outlined in section 3.5.

Figure 4.1 SRN for semantic category 'People'



4.1.1 Identity Measures

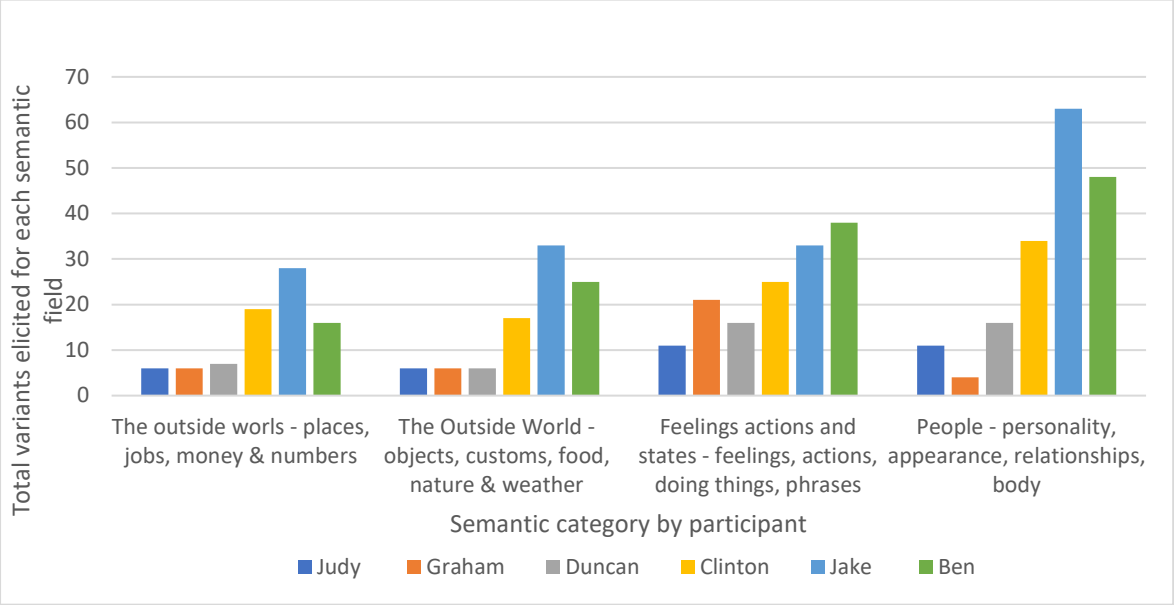
Table 4.0 above depicted the total non-standard variants elicited from each participant using the SuRE method. These figures will be contextualised through the following analysis in relation to each individual's circumstances and influential factors. These figures included variants elicited using the SRNs as well as lexical items elicited during conversations through discussions concerning SRNs. Discussion was also elicited using the Identity Questionnaire outlined in 3.6 which additionally elicited non-standard variants. As outlined in methodology section 3.6, the Identity Questionnaire contained a semi-structured interview questionnaire with questions relating to identity and language. The ID questionnaire can be found in Appendix 2.

Table 4.0 Total number of non-standard variants inclusive of repeated variants across participant across all four SNSs

Participant	Judy	Graham	Duncan	Clinton	Jake	Ben
Total number of variants elicited	38	36	42	97	153	141

The SRNs consisted of 212 English lexical variables used to elicit 'non-standard' variants. A total of 507 variants were elicited from the six informants as part of the lexical data collection. Additional words not on the SRNs produced by participants were also added. Table 4.1 indicated that Clinton, Jake, and Ben produced considerably more non-standard variants than Judy, Graham and Duncan. Observations of difference between participants will be analysed and discussed within the context of the historical and contemporary archive. These data sets are described in section 3.5.3 and referenced in Tables 4.6 –7.7 within the analysis chapters. The characteristic pattern of elicitation is further illustrated below in Figure 4.2. Figure 4.2 illustrated the total lexical items elicited for each SRN for each participant.

Figure 4.2 Item frequency by semantic category for each participant



The semantic categories displayed above revealed those items elicited within each SRN for each participant. On observing the graph, Clinton, Jake, and Ben produced comparably high frequencies of items. Judy, Graham, Duncan and Clinton were interviewed face to face whilst Clinton, Jake, and Ben were interviewed via post and email. As part of the analysis Graham, Clinton, and Duncan possessed full transcriptions of audio recordings. Judy declined the option of recording the interview, and, following ethical procedure, Judy's request was adhered to, with written records and field notes being used instead. Judy agreed to her responses being transcribed during the interview, and her lexical responses were recorded during the interview.

The analysis of data within these chapters, collected through methods described in sections 3.5-3.7, has given emphasis to background data shown in Table 4.1 below. The analysis also considered identity measures shown in the Tables below. The Tables below depicted the self-reported measures from the Identity Score Index in Table 4.2. The identity score had a maximum of 14 and Duncan was rated with the highest There were two scores that represented a positive attitude to those ethnically or regionally close to the individual. In Table 4.2 figures for some participants were not collected for practical reasons. As discussed in section 3.8.2, two of the interviews took place through mail and email correspondence and this raised practical issues (see section 3.8.2) related to time limitation and the need for support to complete all of the research tools. The Relational Analogue Scale is shown in Figure 4.3. Again, only those that were able to complete that element of the questionnaire are shown in Figure 4.3. Table 4.3 revealed the participants' estimation of neighbours, friends and family who identified

as Roma or Traveller. This represented the participants' perceptions of their social network (see discussion section 3.8.1 for limitations). Table 4.4 illustrated self-estimations of social networks that included number of friends and neighbours from within the community. These data are discussed within the analysis and within the framework for analysis outlined in 3.8 and below. For a full description of these research tools see chapter 3 section 3.7.

Table 4.1 Background data

Participant	Judy	Graham	Duncan	Clinton	Jake	Ben
County/Region	Leics.	Derby.	Notts.	Notts.	Lincs.	UK/New Z.
Age	30s	30s	40s	60s	60s	60s
Sex	Female	Male	Male	Male	Male	Male

Table 4.2 Identity Score Index

Participant	Judy	Graham	Duncan	Clinton	Jake	Ben
Regional Index Score	4/14	N/A	6/14	5/14	N/A	N/A
Ethnicity Index Score	4/14	N/A	6/14	5/14	N/A	N/A

Fig 4.3 Relational Analogue Scales

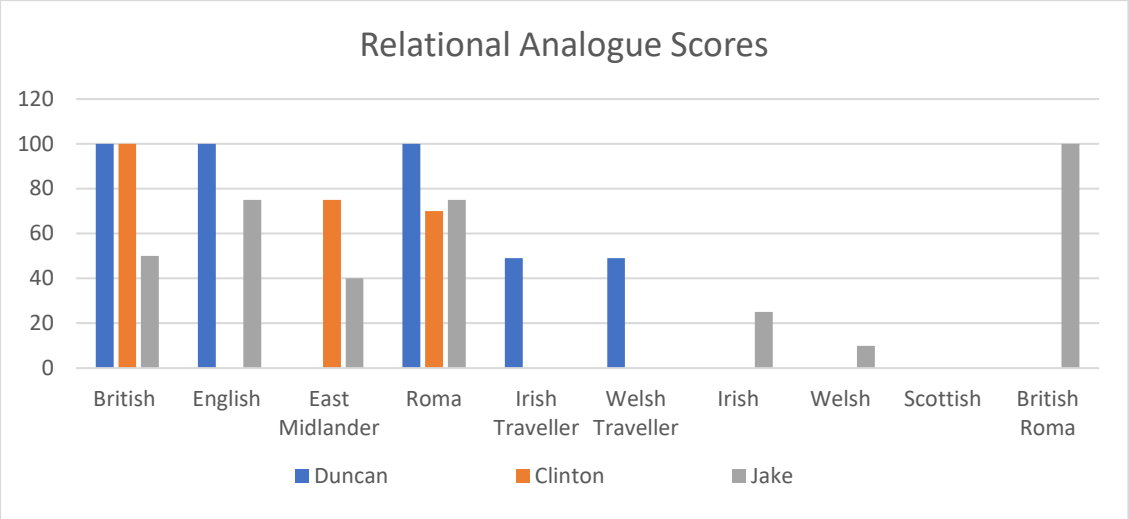


Table 4.3 Social Network

Participant	Judy	Graham	Duncan	Clinton	Jake	Ben
Neighbours with Traveller or Roma background	20%	N/A	5%	0%	20%	N/A
% of friends identity as Roma or Traveller	20%	N/A	20%	20%	40%	N/A

The primary focus of the analysis was to uncover structural variation within the sample of lexical variants taken from the six participants. Variants elicited have been contextualised through comparisons with the archive data sets. Phonological variation will be described where appropriate and limited to where transcriptions were available and comparable. For further discussion of how the distinction between lexical and phonological variation is defined, see section 3.5.3. The study's aims were to uncover, through fine-grained analysis of the patterns of variation, the variables that influenced those variation. The interview data additionally enabled an opportunity to analyse lexical use, together with meta-linguistic narrative and frame this with the principles of indexicality. These principles are outlined in section 3.8 and related to Bucholtz and Hall's (2005) framework.

In addition, the aim of the analysis of the interaction was to uncover how such discussion-work was used as a semiotic tool for linking linguistic form as a social practice to perform identity (Ochs, 1992; Silverstein, 1985). The basis for this analysis further used the framework of the principles of identity proposed by Bucholtz and Hall (2005). This analysis would therefore frame lexical usage, including its meta-linguistic discussion as a tool for creating social meaning and performing social identity. Following is an analysis of lexical variation and of meta-linguistic discussion of Anglo-Romani and Irish Traveller Pavee Cant speakers, predominantly, within the East Midlands. For a description of the six participants, refer to chapter 3, section 3.3.1. The chapters are organised in terms of semantic categories of the SRNs. Within each chapter of analysis, those variables (standard English prompts) that elicited the greatest variation between participants have been discussed, first, with those producing least variation last.

4.2 Semantic category: The Outside World 1

The SRN 'The Outside World 1' is divided into four categories. These categories are 'Food' 'Objects', 'Nature and Weather' and 'Customs'. As with all four SRNs each sematic category related to the central node concept. The development and selection for each central node and related subcategory is outlined in section 3.5. Detail has also been provided in section 3.5 to elaborate on the procedures for development of the SRNs. There were a total of 97 variants recorded for the SRN 'The Outside world 1'. As with all chapters the analysis will highlight evidence of structured variations of lexical retrieval and constructs of identity formed through social interaction.

4.2.1 Food

The semantic subcategory for SRN 'The Outside World 1' that produced the most variants was 'Food'. Lexical variants produced for variables 'food', 'water', 'meat', 'milk' and 'potatoes' were the same for all participants. Table 4.4 shows the responses for the fifteen English variables related to 'Food'. The codes below itemise each variant in Table 4.6 to correspond to archive data sets shown in Table 4.7.

Table 4.4 SRN: The Outside World 1: Food

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Food	*Fodder*	*Hoben/Scran* B2, DV, LSP, N1, SC1, W1/MC		*Oben* MC	*Hoben* B2, DV, LSP, N1, SC1, W1	*Hoben* B2, DV, LSP, N1, SC1, W1
Alcohol		*Peeve (on the Peeve - on the drink)* MC	*Peeve* MC		*Livna*	*Tato pani*
Spirits						*Tato pani*
Beer					*Livna* DV2	*Livanaki/ livinor* /S1
Wine						*Mol* MC

Red Wine						*Loolo mol*
White wine						*Purno mol*
Water		*Parny* IV,D1, LSP, N1, R,G,S1, SC1	*Paani* IV,D1, LSP, N1, R,G,S1, SC1	*Pani* D1, F1, FS, S1, SC1, T1, V2, W1, L1	*Parni/pani* IV,D1,LSP,N1, R,G,S1/ D1, F1, FS, S1, SC1, T1, V2, W1, L1	*Pani* D1, F1, FS, S1, SC1, T1, V2, W1, L1
Cup of Tea				*Pia Mongerer S1, B2, MC*		
Meat	*Faiten*			*Mass* IV, D1, R, N1, R2,G,S3	*Mas* IV, D1, R, N1, R2,G,S3	*Mas* IV, D1, R, N1, R2,G,S3
Milk				*Tud* D1, N1, R, G	*Tud* D1, N1, R, G	*Tud* D1, N1, R, G
Potatoes				*Puvingra/ pulmingerer* N1, S1, B2. MC/	*Puvengra* N1, S1, B2, MC	*Povengrie* N1, S1, B2. MC
Chicken					*Kani* D1,V1,F1,B2,SC1,L1	
Sheep					*Bokra* D1,W1,B2, MG	
Cheese					*Kal* D1,R,G,FS,S3,SC1	

Table 4.5 Archive data sets*

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.

20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

*Each code from table 4.5 (e.g. IV) found in table 4.4 indicated an elicited variant found in the corresponding archive source (see section 3.5.3 for description of archive sources).

Within the subcategory 'Food', in terms of variation, variables 'food' and 'water' produced the most variants for this subcategory which has been recognised as a fundamental of life. The term for 'food' in the Anglo-Romani dialect is found across a number of sources of Anglo-Romani lexis. These date from across 19th and 20th Century sourced vocabulary lists. The items *scran* and the word *oben* /ɑbən/ with a dropped /h/ (not *habən* /) is recorded more recently as part of the Manchester Corpus (2021) compiled over recent decades. *Scran* is attested as a Cant term by the Manchester Corpus. This aligned with Duncan's use and identity as he described himself as having both Roma and Irish Traveller parents, and explained the impact this has had on his childhood, culturally in terms of language and family life.

Judy identified as 'Irish Traveller' and used the term 'Pavee Cant' to refer to the lect or variety associated with this ethnicity. For Judy, two variants *fodder* and *faiten* for 'food' and 'meat' were elicited from within this semantic category. *Fodder* is not attested as a dialect variant within the archive sources although the term *fodder* is attested as an old English term, rooted in proto-German (Kroonen, 2009) defined as bulk food for cattle. Several of Judy's responses were not found within the archived sources categorised as Irish Pavee Cant or Anglo-Romani. This was indicative of a change in progress with respect to the literature surveyed within the archived sources (Table 4.5). Judy's dialect repertoire mixed non-standard English with vernacular (i.e. use of *fodder*). The process of lexification of ethnic dialect into non-standard vernacular English within more mainstream culture is analysed and discussed within the

proceeding chapters. The process of the integration of slang or non-standard lexis into the ethnic dialects of Irish Traveller and Roma is also analysed and discussed. Within the following chapters, this is explored as a characteristic of Judy's dialect repertoire as she often used vernacular English within her dialect repertoire. Conversely, there were examples of Romani and Cant words used by non-Roma and non-Irish Traveller participants. This is understood to be part of a wider pattern of use, where use of such terms invoked an affected stance (Bucholtz, 2012), such as, local solidarity between non-GRT identities, and resembled an enaction of kinship, acting as an affective marker of affiliation that was displayed by those who identified as part of a GRT community or was distinguished as someone with a related heritage.

Within this SRN, Graham, Duncan, Clinton and Ben all produced the variant *hoben* with Duncan dropping the initial /h/ realisation in /ɒbən/. In excerpt 4.1 below Graham was discussing the use of the dialect together with a local close friend (see section 3.1.1 for further description of participants in the interview context). Here they were discussing the use of the variant *scran* for food.

Excerpt 4.1

1	Graham:	*you'd only use it like among Gypsies wouldn't you.*
2	Acquaintance:	*that's right yes*
3	Graham:	*you wouldn't use it*
4	Acquaintance:	*you wouldn't use it outside*
5	Graham:	*no no I mean no one would really understand some people now*
6		*they'd say gel the gel and get some scran… but that's a bit more*
7		*modern now*

Most of the discussions that took place throughout were framed around the meta-linguistic discussion topic of word use and form. Graham in line 1, used modality and a question tag with falling intonation to reflect a tentative, evaluative stance on his proposition that would be conducive (Hudson, 1975) to the response that '*you*' would only use certain words amongst 'Gypsies'. Graham went on to is explain the terms *gel* (later discussed in section 6.1.1 SRN: Feelings, Actions and States: Doing Things) and *scran*. Here, Graham contrasted the use amongst 'Gypsies' with modern use of Romani lexis by stating 'that's a bit more modern now' with the implication that people outside of the community might use the term *scran*. Graham,

in line 5, marked modality using '*a bit*' to propose a negotiable epistemic stance on the use of Romani lexis as being '*more modern*'. The implication being that the term '*scran*' was no longer indexical to an identity that participants within this study related to, whereas, used embedded within a phrase structure, such as, *gel the gel and get some scran* would be indexical to identity, and framed here within a meta-linguistic discussion. It is argued that background data suggests Graham identified himself as a peripheral or marginal community member (see section 3.5), however, the stance taken in this and in other extracts (to follow) suggested a positive attitudinal positioning of identity adopted also by other participants within the study, notwithstanding the limitations of the attitudinal measures discussed in the methodology section 3.8.2. As with Judy's inclusion of English informal vernacular , other examples of Anglo-Romani and Irish Traveller dialect have been assimilated into vernacular status. In the example 4.1 in line 5, the evaluation of the word '*scran'* as: '*that's a bit more modern now'* implicated a change in progress of that lexical item. There was an indirect connotation of ownership and indexicality of lexis and identity and the changing characteristic of words such as '*scran*' implied a repositioning (Bucholtz and Hall, 2005) over time. One conclusion might be that some words were no longer part of community identity and were repositioned within social use. Further examples explored this phenomenon, focusing on words that were discussed in terms of their relation to participants' self-ascribed identities, and analysed from discussions around meta-socio-pragmatic use.

Excerpt 4.2 from Clinton offered an example of use of the Anglo-Romani term *hoben* with explanation and a translation into English from Anglo-Romani. There was an implication that the Romani term was not used in the same way as the standard English word 'food' thus indicating an extension of usage.

Excerpt 4.2

1 Clinton: *you'd just say **oben** stuff **oben** means vegetables it means a lot of things*

The above utterance was an example of Anglo-Romani lexis used as a superordinate to reference a number of concepts such as '*vegetables',* indicated with *'it means a lot of things'*. This superordination is observed across the data for a number of variants and is highlighted in examples through the analysis chapters. Whilst the use of second person plural pronoun '*you'd just say **oben** stuff*' suggested no specific identity work, extract 4.3 below cited the use of the exonym *Gorje*, the referent term used for outsiders, used to refer to those outside the

community and indicated social positioning through lexical choice (this is discussed in detail in chapter 7.4, 'People, Age and Relationships')

The term for 'water' in Table 4.6 revealed it was consistent across Graham, Duncan, Clinton, Jake and Ben with minor variation in terms of spelling. Whilst spelling was reported problematic by two of the participants, it was common for spelling to be inconsistent across dialect sources. Whilst literacy might be considered as a factor accounting for variability amongst participants, spelling variation within and between participants' dialects were considered less significant although an integral (Beal and Cooper, 2015: 33) but often overlooked aspect of dialect variation. Whilst significant patterns of phonological distinction were not discernible (see section 3.5.3 for detailed lexical comparisons). In the following excerpt Duncan was explaining the use of the word for 'non-Traveller' as an aside whilst explaining the term for 'water'.

Excerpt 4.3

1 Duncan ***Gorje*** *which is a non-Traveller and that's as I say that's used to this*

2 *day that's used very common very very common every day you would*

3 *say that word its being used you know what I mean because we're*

4 *meeting you know different people and I would say I spoke to* ***Gorje***

5 *man today you know and that's a word I would use amongst the family*

6 *my children would say that very very often you know and eh a glass of*

7 *water a glass of* ***parny***

This aside indicated positioning between the speaker and the interviewer with the self-declaration of the referent *Gorje*, which implied an indexical association with community identity. This is a good example of an address term (AT) (Kiesling 2012: 177) and one that Duncan chose to use with the cultural implication of stance taken within a GRT identity. The use of the term *parny* also implied an emergent stance (Bucholtz and Hall, 2005; Bucholtz, 2012) with use of the variant within the context of the meta-talk concerning lexical usage. This meta-linguistic talk was further explored through the analysis of interview to explore the social meaning of lexis, the stance taking or positioning from within the interaction themselves, with an understanding that speakers ascribed to their use of this lexis.

The term for ‘water’ elicited from the participants of this study was consistent with archive sources, whilst there appeared to be a degree of phonological distinction. The distinction between long and short initial vowel realisation was represented orthographically with a double vowel *paani* and ‘r’ *parny* (pɑ::ni) for long vowel, and a single vowel *parni* (pɑ:ni) for a short vowel realisation. This can be considered a regional characteristic with a number of other examples within these chapters.

In the following excerpt participant Graham offered a different stance in explaining the variant for ‘water’.

Excerpt 4.4

1 Graham*:* *water the horse so **parny** is water which is like an Indian word isn't it*

2 Acquaintance: *yes it is yes*

3 Graham*:* *they still use that uh **parny** as water*

In this excerpt, Graham created a distance through interactional meaning with use of ‘they’. The use of the third person plural pronoun alongside the use of a tag question for confirmation of propositional content (*parny* as an Indian word) positioned Graham within the interaction as an observer and commentator. Although this created distance between himself and lexical use within this interaction this should be considered in relation to the context of the interview circumstance. Graham ascribed himself as part of a Anglo-Roma culture and was positioning himself as observer during the interview and could be viewed as social positioning. This suggested an openly nuanced identity within his identity practice. The use of pronouns as an identity practice is referred to again later within the analysis.

'Meat’ was consistent with archive sources in terms of orthography. The term for ‘meat’ was uniform between the Roma participants Clinton, Jake and Ben. All three produced the term *mas(s)*. Excerpt 4.5 showed Clinton explaining *mas* used as a superordinate term for liver.

Excerpt 4.5

1 Interviewer: *meat*

2 Clinton: ***mass***

3 Interviewer: *so like liver would be mass*

4 Clinton: *yeah yeah*

The response in excerpt 4.5 indicated the superordinate term was retained in contrast to the subordinate specific lexical items and was a characteristic of Anglo-Romani within this analysis. This was consistent with corpus data from the archive as a general pattern.

The term for 'milk' is given as *tud* , which was also in line with previous attestations as shown in Table 4.6. 'Milk' elicited *tud* from Clinton, Jake and Ben. Excerpt 4.6 showed Clinton emphasising the generality (as in excerpt 4.2) of the use of *hoben* whilst producing *tud* for 'milk'.

Excerpt 4.6

1 Clinton: *just most of it's* ***oben*** *milk would be* ***tud***

The word for 'potato(es)' varied in orthography (*Puvingra, Puvengra, Povengrie*). However, instances where contrastive phonological realisations were inconclusive were not distinguished for the purposes of this analysis (this approach is outlined in section (3.5.3). Therefore, these variants have been classified as comparative for the purposes of this analysis. The variants were all comparable with four of the archive sources. Comparable data included Clinton, Jake and Ben's productions of similar variants for 'potatoes'. By comparing *puvingra* (/pu:viŋgra/) and *puvengra* (/pu:vəŋgra/) provided limited evidence of phonological distinction with the second syllable pre-consonantal short vowel /i/ compared with orthographic representation of the close-mid front vowel /e/ indicating a contrasting realisation. Ben provided the ending *-grie* contrastive to *-gra*. This distinction was consistent with evidence from the archive sources of linguistic inflection that varied between participant.

One of the single variant elicitations is *pulmingerer* for 'potato'. Excerpt 4.7 showe Clinton's response to 'potato' and highlighted this as significant to region, community and food culture as this is one of few subordinate terms used for a specific vegetable.

Excerpt 4.7

1 Clinton: *well potatoes* ***pulmingerer***

2 Interviewer: ***pulmingerer*** *okay swede*

3 Clinton: *no no… you'd just say* ***oben*** *stuff* ***oben*** *means vegetables it means a lot*

4 *of things*

In Excerpt 4.7, the unreal conditional '*you'd just say…'* within this meta-talk Clinton has placed himself as a speaker of Anglo-Romani. His authority of knowledge or epistemic stance was evident in the referral to the lexical variant '… *oben'* as the implication 'if you were speaking Anglo-Romani…' orientated himself as the expert or knowledgeable person within the exchange. This positioned himself as belonging to a group of Anglo-Romani speakers and to this extent as an authority of the appropriate use (Holmes and Meyerhoff, 1999).

Both Jake and Ben respond with variants for 'beer'. Jake produced *livna* and Ben produced two variants *livanaki* and *livinor*. *Livanaki* for 'beer', which were not found amongst Anglo-Romani sources. This might represent a variant that derived from a European Romani dialect, and, consequently, represented an international Romani influence. Ben identified as Roma within an international community of Roma speakers. His experience and language repertoire represented an opportunity to compare and contrast linguistic practices of those from the East Midlands and a speaker from elsewhere. His inclusion enabled influences of lexical variation relating to historical exposure, and from where opportunities for practice would be significantly different. Ben's variety could be considered in the context of his background (see section 3.3.1) and multi-word responses, including the degree to which lexis for Ben retained some inflectional features within his Anglo-Romani repertoire as significant. Other examples have been discussed with analysis of origin and innovation where available through chapters 4 to 7. The European Romani word for 'beer' is recorded as *lovina* (Manchester Project, 2021). The word is considered to be derived from Slavic origin. The word *lauena* was transcribed by Andrew Boorde (1547) in a first written account of Romani in the British Isles and considered Slavic in terms of Romani origin. As discussed in the introductory section 2.2.1.5, and section 2.3 this account was considered an early record of Romani in England and therefore a point of origin for Anglo-Romani. The Slavic influence spread across northern Romani dialects across North Russia, Latvia, Lithuania, Slovakia and Finland, as well as appearing in Anglo-Romani and Welsh inflected Romani (Matras, 2012). *Livna* was elicited from Jake which corresponded with Dawson's 2002 lists collected in the Derbyshire East Midlands region. This suggested that *livna* (/livnə/) was a regional variant in terms of phonology as all other variants being represented with a vowel post voiced labiodental fricative. There was one exception to *Liv'no* (/livnəʊ/) listed by Sanderson (19thC) although the final vowel is represented as being closed. There was use of Anglo-Romani terms used for more than one English variable as an extension, as Jake used *livna* for both beer and alcohol and Jake reported using *tato pani* for 'alcohol' and

'spirits'. This indicated a tendency to extend the use of lexis as a strategy for maintaining dialect, as specified with other examples from this study.

The remaining variants were single elicitations from individual participants. Jake and Ben produced ten and twelve variants respectively and were, in addition to those mentioned in the SNS. Their degree of productivity corresponded to their background, whilst methodological differences involved in data collection were also considered as interviews were conducted in-person or by email. Clinton produced responses during interview rather than through email correspondence. It could be hypothesised how pressure to perform could affect spontaneity and have been influential towards the numbers of lexical variants recorded (see limitations section 3.8.2) . Clinton producing a relatively high number of variants, although fewer than Jake and Ben.

Single productions of variants included the lexical item *pia mongerer* meaning 'a cup of tea'. This term was found in 19th Century lists, and marked a continuation of this Anglo-Romani dialect variant. The term for 'cup of tea' was also recorded as *muttermongri* or *pi a muttermongri* in the Manchester Corpus, and translated as 'urine tea'. The word pia is attested as old Welsh Romani (Sampson, 1875). Excerpt 4.8 continued to show Clinton explaining the use of the dialect variant for 'cup of tea'.

Excerpt 4.8

1 Clinton: ***pia mongerer***

2 Interviewer: what's *mongerer*

3 Clinton: *it's just [a] word and that what it means a cup of tea*

In this instance, Clinton indicated the term ***'pia mongerer'*** was not considered a compound and he replied: ''*it's just [a]* word'. This suggested for Clinton at least, lexical phrases or multiword phrasings could be fixed, and not able to be parsed, evident from this example. This was indicated not to be the case for Jake and Ben with examples of inflectional morphology evident within their use discussed through chapters 4 - 7.

In excerpt 4.9, Clinton, in relation to this lexis, created an aside about learning languages, and his granddaughters learning French as seen below. Roma and French have been compared in status. Clinton compared the learning of Roma with the learning of a second language in this case the learning of French, and his granddaughter's learning of the Anglo-Romani dialect.

Excerpt 4.9

1 Clinton: *me granddaughter's been learning French and I says to her well I can talk*

2 *French can you grandad yeah parle vour ...**pia mongerer grave**..that ain't*

3 *French that ain't French*

In this narrative, Clinton related the acquisition of Anglo-Romani to the learning of a second language, as his granddaughter was learning French and he stated within the same conversation how he wanted her to learn Anglo-Romani. Clintons exclaimed '***pia mongerer grave**..that ain't French'* to compare the learning of French with Anglo-Romani to comical effect. Clinton contrasted learning a second-language in school with learning Romani within his own family. The appraisal of Roma as comparable to the learning of French raised the idea of the value of Roma. This question or uncertainty was characterised within the narrative that took a stance of countering an assumption that the Roma language was of less educational value. This interaction showed Clinton relating his own identity as an Anglo-Romani grandfather positioned within a narrative framing the value of learning Roma. The interaction was relational (Bucholtz and Hall, 2005) in that Clinton indexed his Anglo-Romani identity through this narrative.

In terms of productivity of lexical variation, Table 4.6 revealed several variants that were produced by single participants most notably by Jake who produced variants in relation to the central node concept of food. Jake has discussed his work promoting Roma culture and this suggested a wider social network (see limitations section 3.8.2), and was, therefore, actively promoting Anglo-Romani culture and language which was a factor relating to higher productivity. These were variants for the variables: 'chicken' *kani* , 'sheep' *bokra* and 'cheese' *kal*. Ben also produced variants for 'wine' *mol*, 'white wine' *purno mol* and 'red wine' *loolo mol,* as well as, 'alcohol' attested as synonymous with 'spirits' *tato pani*. As mentioned, Jake and Ben were interviewed via email (see section 3.8.2 for discussion of methodology) which might have influenced productivity. This was further discussed as a limitation in terms of data collection. However, background factors have been argued to show a more significant and influential factor for their degree of productivity and evidently in the case of Jake. Jake was the only participant who identified as British Romany (see Figure 4.3) and perceived his social network (see limitations section 3.8.2) as more significantly as having a Roma identity (see Table 4.3). Jake also produced *loola mol* and *purno mol* for red and white wine respectively. Although these were not attested in the archival sources for Anglo-Romani it was likely a

common collocation, innovation or characteristic of Jake's morpho-syntactic productivity, as *lolo* was European Romani for red and *parno* for 'white', whilst wine was *mol*. *Mol* was an Anglo-Romani variant with several orthographic and speculative phonological variants.

The variants for 'chicken', 'sheep' and 'cheese' have been found across a number of sources. Phonologically similar but distinct variants have been found within the archive. The term *kani* from Jake in European Romani has been recorded as *khajin* meaning 'hen' and attested as a root form (Manchester Project, 2021). The variants for 'sheep' and 'cheese' have been similarly found across archival sources with orthographically comparable but distinct variants across several historical lists and sources. The term for 'sheep' was recorded as '*bakro*' in European Romani and 'cheese' as '*kiral*' in European Romani (Manchester Project, 2021). These words demonstrated semantic areas that were linked to rural and agricultural communities. Jake was from Lincolnshire, and this was potentially an association that was linked to his background, with a potential factor for lexical maintenance within a rural and agricultural semantic field (see 3.3.1 for participant description). Judy produced the term *faiten*, which has not been recorded within the literary sources for Cant or Romani although *fe* has been recorded by Rieder (2018) for meat, and seemed likely a related term.

4.2.2 Objects

Table 4.8 depicted overall patterns of productivity, comparable to previous semantic categories. Judy, Graham and Duncan produced comparably fewer responses than Clinton, Jake and Ben (See Table 4.0 for comparison across all SRN). As mentioned in section 4.2.1, participants Jake and Ben were interviewed via email, considered as a methodological limitation (section 3.8.2), which might have influenced responses owing to factors of pressure and opportunities to consider a response. An ethical reflexive approach using semi-structured interviews (see section 3.2) provided advantages in terms of authenticity (Skinner, 2012). However, there were limitations to this approach and its application in this study (discussed in section 3.8.2) were the results of constraints of email as a means for conducting a semi-structured interview. Participants Judy, Graham, Duncan and Clinton were interviewed and relied on their ability to produce lexis spontaneously. With reference to the results of the SRN below, Table 4.8 showed Judy produced 3, Graham produced no responses, Duncan produced 2, Clinton produced 6, Jake produced 12, and Ben responded with 6 variants.

Table 4.6: SRN: The Outside World 1: Objects

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
This				*Kuvver* R, V1, F1, SC1, (R)	*Dova* D1, L1	*Aka/akla*
Thing	*Youk* DV4			*Kuvver* Ev,SC1,L1	*Kie/Koava* /D1,	*Kova* Ev,SC1,L1
Car			*Rog/Screeve* DV4/MC			
Door	*Inik*			*Jigger* MC,DV4	*Wooda,/Vudda* MC,IV,D1,R2,S3, LSP,F1,B2,SC1,L1/	*Wuda* MC,IV,D1 R2,S3,LSP F1,B2,SC1, L1
Soap				*Sappin* R2,MC,S1	*Suprin/Sapna* D1,DV1,DV3/MC,L1	*Tovengrie* (wash)
Stick					*Kosh* MC,LSP	
Firewood					*Yogkosh* DV1	
Book					*Lil* R,N1,R2,G,FS,SC1,S3,T1,F1,S1,W1,B2,,SC1	
Poem				*Sorla*	*Gili* DV1,W1,B2,SC1, S2	
Kettle	*Saucepan*			*Parni kova* B2,W1,SC1,	*Kekaari* DV1	*Kekavaki*
Bicycle				*Proste*		

Table 4.7 Archival data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

In terms of variation between participants, the variants for the English demonstrative 'this' were contrastive between Clinton, Jake and Ben. As a function word it showed there were grammatical characteristics that differed between the participants. This was a distinguishing feature of lexical repertoire and individual variation uncovered through this analysis. Table 4.8 demonstrated both *kuvver* and *dova* were drawn from Duncan and Clinton, both of which appeared in the archival data sets. The word *kuvver* has been recorded in several 19th century sources, and was a cognate form of *(a)kava*, a European Romani word meaning 'this'. Jake used *dova*, which was recorded in the more recent vocabulary of Dawson (2002), a 'folk dictionary' (see section 3.8.2) although this was featured within the Manchester Corpus (2021) as well as in the earlier Lucas (20thC). *Dova*'s etymology was located with the European Romani demonstrative *(a)dova* which meant 'that'. It appeared the distinction between the two concepts might have been assimilated within the usage of Anglo-Romani, at least for participant Jake with both meanings attached to the word *kuvver* and indicated a change in progress in relation to Anglo-Romani sources (Matras, 2010). In excerpt 4.1.1 , Clinton explained the usage of the term *kuvver:*

Excerpt 4.1.1

1 Clinton: *you'd refer to something if you was looking at something or someone's got*

2 *something where* ***dick*** *to the* ***kuvver akai*** *that means look at the thing you know*

3 *what I mean*

The word *akai* was also used in except 4.1.1 The word *akai* was also recorded as used for 'this' in archive sources (Manchester Project, 2021). The meaning of the word used in this excerpt, however, conveyed the meaning 'here'. The word *akai* was recorded as European Romani meaning 'here' (2021). It was not clear from this example whether *akai* and *kuvver* were interchangeable in certain contexts of use or if they had been assimilated for some Anglo-Romani speakers. For Clinton however, these wee two separate lexical items as illustrated in excerpt 4.1.1

Jake produced *dova*, which was also European Romani *(a)dova*, for 'this'. This suggests that there was assimilation for this participant for those demonstratives. Archival data (Manchester Project, 2021; Dawson, 2002) would further suggest this to be the case more widely, whereas a general pattern has been described as an inherited complex system for Welsh and Anglo-Romani overall by Matras (2010a). Ben produced *aka* and *akla*, both of which did not appear as orthographic matches in the archive's sources. The closest orthographic approximation was *akaj,* the European Romani word meaning 'here'. This suggested a common etymology. Again, Ben produced several distinct variants compared with the other participants that self-identified as Roma or Irish Traveller. These were indicated in section 4.1.1, with variations between participants suggesting a nuanced and complex picture for each participant's identity. There was a clear distinction in terms of Ben's overall dialect variation, which linked to several of his background factors. One notable factor was his connection to an international Romani network (see section 3.3) which could be expected to influence his variety of an Anglo-Romani dialect. Those factors are further considered through examples explored throughout these chapters.

The variable 'thing' produced uniformity across participants. There existed some comparability between *kover*, *koava*, and *kova* for Clinton, Jake, and Ben, respectively. The term *kuvver* for Clinton was translatable as both 'that' and 'thing' within the context of the interview as discussed in reference to excerpt 5.0. It might be assumed that *kuvver* could be used for both 'this' and 'thing' although this would need further clarification. Notably, there was less variation within the archive sources for 'thing' appearing as approximations of *kuvver*. Jake produced *dova* being a cognate of *(a)dova* (European Romani 'that'). Clinton, Jake and

Ben all produced variants that were cognates of *(a)kova* (European Romani 'this'). The variants for 'thing' all shared the same root *(a)kova* meaning 'that' in European Romani, suggesting some consistency in terms of the variants associated with that, and used for 'thing'. This compared with the distinction for 'that' and 'this' where there was wider variation.

Judy produces *youk* for 'thing' which was found as an approximation of variants with the same meaning within the archive's sources. The approximations were: *aynokh, enokh,* and *inokh* (Dawson, 2011*)*. These demonstrated a degree of comparability with Dawson's (2011) vocabulary list (see limitations section 3.8.2) elicited from those that self-identified as Irish Traveller Pavee Cant speakers. Judy self-ascribed as Irish Traveller and perceived her local residents and friends identifying as approximately 20% Roma or Traveller. Whilst spontaneous recall under interview constraints might have affected elicitation, this elicitation indicated a maintenance of this lexical item.

Other comparable variants were those obtained in response for 'soap'. Both Clinton and Jake respond with comparable variants *sappin* for Clinton and *suprin* and *sapna* for Jake. The Romani word is recorded as *saponi* (Manchester Project, 2021) meaning soap in European Romani. Excerpt 4.1.2 illustrated Clinton's use of the word *sappin*.

Excerpt 4.1.2

1 Clinton*:* *because you'll say if I think he ain't had a wash today I'd say* ***dick*** *to the* ***gearer***

2 *he had no* ***sappen*** *this morning that means he hasn't had a wash this morning*

3 *are you wi'me* ***sappen***

In this excerpt Clinton narrated a hypothetical anecdote to illustrate the use of the word *sappin*. There were several references to cleanliness within the interview data, and was a subject matter that continued through several exchanges with the interview data. Cleanliness was mentioned as conceptually significant within Romani culture and discussed within sources dealing specifically with cultural concepts, such as, the need to dispose of chipped or broken crockery (Okely, 1983). The story in 4.1.2 is presented as a humorous illustration of the meaning and use of the word *sapan* with reference to the cultural context of highlighting awareness of cleanliness, and, in so doing, implied an evaluative orientation towards being unclean in: '*had no* ***sappen*** *this morning'.* In this example, Clinton offered an affective evaluation of an incident of uncleanliness, which indexed an ideological position of the value of cleanliness. Table 4.8 also showed Jake produced two variants for 'soap', both of which are found within Anglo-

Romani word list archive sources. However, Ben produced an innovation in *tovengrie*. This was a derivation of the European Romano verb *thov* 'to wash' with the nominalisation -*engrie* affixation adding the meaning of pertaining to or belonging to an object and appeared to be outside of the current mention within the literature of Anglo-Romani.

The terms for 'kettle' were also comparable. Jake and Ben produced *kakaari*, to b\e found in the Anglo-Romani archive sources, and *kekavaki* which was not found in the word lists. This followed the pattern of variant production with respect to Ben being more productive in terms of inflection and innovative forms (Table 4.2), characteristic, perhaps, of being raised by a Romani speaking father (section 3.3.1) and his self-described proactive involvement in Romani cultural promotion. However, with this example, the term used by Ben was closer to the European Romani form *kekavi* or 'kettle', although all three were comparable and categorizable as cognates. Clinton produced *parnokova* which as a compound noun did not feature in the Anglo-Romani archive sources. This word could be a new attestation and would need further investigation to confirm.

There were several single response variants recorded for variables within the subcategory Objects. 'Car' evoked *rog* and *screeve* for Duncan. *Rog* is recorded as Cant (Dawson, 2011), orthographically recorded as *raug* which was a derivation of *roglan* meaning four wheeled cart (Dawson, 2011). The variant *screeve* is also attested as Cant lexis (Manchester Project, 2021). Duncan self-identified as Irish traveller, Welsh Traveller as well as Roma, British and English (see Table 4.3). He associated his use of these terms as part of his ethnic inheritance from both his mother and father whom he described as Irish and Roma. Jake produced a number of variants outside those elicited from within this category of SRN objects. *Kosh* for stick, a term borrowed into English from Anglo-Romani is given. The European Romani root term is recorded as *kašt* (Manchester Project, 2021). The variant compound noun *yogkosh* also produced by Jake was also found in the archive source. The term *yog* or *yag* meant fire in European Romani. *Lil* is produced by Jake for 'book' was featured across several Anglo-Romani vocabulary lists, whilst Duncan produced the variant *sorla* for poem when responding to the prompt 'poem'.

In excerpt 4.1.3, Clinton explained that many of the words had the same meaning following on from a discussion concerning the use of different words for 'caravan'.

Excerpt 4.1.3

1 *Clinton: I mean yeah but there's millions of words what you can use*

2 *and thousands of them can mean the same thing you know what I mean*

3 *Int:* *a song a good poem or*

4 Clinton*:* *it's gone (1.5)* ***sorler***

The discussion continued about words for 'song' or 'poem'. Clinton suggested he did know a term by stating 'it's gone' in excerpt 4.1.3 possibly indicating issues with recollection that related to frequency of the use of the terms. He then produced *sorler* after a short pause of 1.5 seconds. This term was not recorded as 'song' amongst the archive sources, and no instance of a term for 'poem' has been recorded. This might be a cognate of *sorla*, which was Anglo-Romani for 'morning', referring to the morning chorus or song of birds, but as there were no comparable sources this was speculative. Clinton also produced *proste* for bicycle which was also a term that could represent a new attestations. *Sasta grai* was recorded in the Manchester Corpus (2021) for bicycle. The term *sastri* was European Romani for 'iron', which indicated as a word that had changed through time to the term *proste* (Manchester Corpus, 2021). *Proste* might be a derivational innovation that had been derived from the Romani term *sastri*. Excerpt 4.1.4 showed Clinton illustrating the use of the term *proste.*

Excerpt 4.1.4

1 Clinton*:* ***proste*** *whats* ***proste*** *a bike*

2 *Int.* *I've not heard that one I don't know that one proste*

3 Clinton*:* ***proste***

4 *Int.* *I mean Newark slang we used to say grid a grid*

5 Clinton*:* *grid does mean a bike thats English slang though isn't it...*

6 *Int.* *how would you use that sort of..get on your* ***proste***

7 Clinton*:* ***chore*** *that* ***proste*** *look* ***chore*** *that* ***proste*** *look*

8 *Sp2.* *that's where they've all gone the police are*

9 Clinton*:* *no no I'm just saying*

In excerpt 4.1.4 Clinton differentiated between terms used in English those he defined as 'slang' and those which were Romani. From this excerpt the two were determined as not interchangeable with marked distinctions as Clinton stated: '*grid does mean a bike that's*

English slang though isn't it.... The explicit mention of this categorisation illuminated the distinction between the status of Clinton's of what he defined as 'slang' and the dialect of Anglo-Romani. In this example, Clinton contextualised the term in the statement '*chore that proste look'* or 'steal that bike look'. The contextualisation of the term *proste* within this phrase indicated a light hearted and comedic use of the word. This enacted a playfulness that contrasted with the less playful task of collecting lexical data that was the actual research task. This attitudinal stance was evident within this interaction and was an attempt to create humour, and, therefore, a more characterful portrayal of Clinton's variety and more, importantly, his Anglo-Romani.

4.2.3 Nature and Weather

The SRN subcategory for Outside World 1, 'Nature and Weather' revealed several variants as distinct between the participants with a degree of comparability between Anglo-Romani vocabulary sources from the 19th century to recent archival sources. In terms of productivity, both Jake and Ben produced eight and nine respectively compared with three, two , one and one for Clinton, Duncan, Graham, and Judy. This followed the pattern of productivity observed for across categories.

Table 4.8 SRN: The Outside World 1: Nature and Weather

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Hay					*Kas* DV2, R2, N1, G, D1, F1, S1, B2, SC1, L1	
Hedgehog				*Hotchy* MC, DV2, W1	*Hochiwichi ochi* MC, DV1,R, T1, B2, /MC, DV2, W1	*Hotchi wa/Urkos/Tikni bawlo* /_/DV1

Horseshoe game			*Coits*			
Horse	*A Filly/A stallion/A Mare*	*Grai* MG, IV, DV2, R, N1, R2, G, EA, LSP, DV2, V1, W1, B2, DV1	*Grie* S3, N1, F1	*A grie* S3, N1, F1	*Grei* MC, T1, SC1	*Grai* MG, IV, DV1, R, N1, R2, G, EA, LSP, DV2, V1, W1, B2,DV1
Mare						*Grasni* DV2, F1, W1, SC1
Foal						*Tika grai*
Sun					*Kam* W3, DV2, N1, R, G, S3, F1, S1, V2, W1, B2, SC1	
Rain					*Brushindo* GY, S1, SC1	*Bresheniskri*
Wind					*Baval* IV, N1, G, GY	
Earth					*Ful* DV2	
Fire				*Yog*	*Yog* DV2, N1, R, R2, G, S3	*Yog/Yag* DV2, N1, R, R2, G, S3/MG, IV, EA

Table 4.9 Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

Variants for 'horse' were elicited from all participants. This was the only variable within the category 'Nature and Weather' that elicited variants from all participants. This was illustrated as a salient term for some speakers of Pavee Cant and Anglo-Romani. This pattern, therefore, could be related for these participants to the prominence of horse culture and does persistthrough several discussions within the interview data. Horse culture was raised within several exchanges and described below. Horse culture was also discussed within certain fields of literature that related to ethnicity of Roma and Traveller related cultural activities (Okely, 1983), where images and associated information were shown onvarious website publications and cultural sources, such as, 'Patrin' Gypsy Heritage in the East Midlands (Duffy et al., 2014). Locally organised horse racing has also been mentioned as a cultural activity. Okely (1983) explained how the horse has been a mediator between cultures and communities, with the principal high value status, which horse ownership received amongst certain community groups. An example below was an excerpt, in which Clinton explained his stance on culture and ethnicity, emphasising tolerance and acceptance within the context of horses a part of his cultural identity.

Excerpt 4.1.5

1 Clinton: *I'm still using horse and carts up to the last eight or nine years tens years ago*

*2 I had eh, they're called **drays** four wheeled **drays** and I was using them here*

3 and there for different bits and bobs and we used to go on the antique fair

4 there we used to take stuff there, my kids used to take stuff from one spot to

5 another where that had containers and that, and we was using them that way

6 you know what I mean,but we was still using a horse and cart for driving down

7 the road

Clinton's narrative mentioned horses and horse '*I'm still using horse and carts up to the last eight or nine years tens years ago'*. This association was foregrounded with a narrative on cultural diversity and his own position as someone who would: '…*mix in with people, where they from won't make no difference…'.*

Excerpt 4.1.5 (continued)

1 Clinton: *you're not ashamed, I'll mix in with people, where they from won't make no*

2 difference. I'll talk to them about one thing or another or they'll say oh you come

3 from so and so yeah, that's right and talk [to] them and they accept it but some

4 of them don't you know what I mean, but you are wherever you're born from

5 you are you know what I mean if you're born a 'fucking' dog you're going to be

6 a 'fucking' dog ain't you or whatever you know what I mean whatever sort of dog you are you are and that's it you know what I mean you're still that but I mean I like to carry on.

This demonstrated as an appraisal of Clinton's own stance on an acceptance of others' differences. His mention of: '*I'm still using horse and carts up'* was partly an evaluation of the importance of the use of the horse and cart, and his mention of his use of: '*they're called **drays** four wheeled **drays'*** has been highlighted here as relevant to Romani culture and his own position within the culture he described, and was evidently strongly associated with his own self image.

In terms of variation, both Graham and Ben produced the variant *grai.* Although many of the variants produced by Ben appeared separate from the vocabulary of the Anglo-Romani word

lists there were those that were comparable and in this case identical. Duncan and Clinton's variant *grie* might indicate a regional pattern as both were from Nottinghamshire (see Table 4.1) Jake's *grei* was distinct in vowel quality as in *grie* /gɹi:/ v *grei* /gɹei/. This revealed a greater degree of uniformity compared with other variants comparable between participants. Judy produced three variants which were not found within sources collected from Irish Traveller Pavee Cant speaker sources, although the inclusion of these terms could indicate horses as culturally emblematic for some speakers of these varieties,

Another variable within this subcategory that produced variation was 'hedgehog'. There were several mentions of hedgehogs within the interview discussions, and, generally, these related to either naming or the practice of eating hedgehogs. In terms of lexical variation, terms for hedgehog are mentioned across a number of literary sources (Matras, 2010). Whilst popular accounts associated Romani culture with the eating of hedgehogs (Okely, 1983), this was anecdotal and not relevant to the focus of the lexical variation, includingthe social meaning of Anglo-Romani in use. However, an anecdote did arise within the semi-structured interviews ,and contextualisation proved a useful measure, as an understanding of how a comical image of hedgehog eating could be made was acknowledged in the following exchange 4.1.6. Such portrayals are cited in literary accounts, such as, in Michael Morpugo's (2006) *Mr Nobody's eyes* where a boy runs away from home and hides with a family of Roma where the young girl hides hedgehogs as she doesn't want her uncle to eat them. Excerpt 4.1.6 was an example of conversation around the topic of hedgehogs and Anglo-Romani culture that related to preparing hedgehogs as food.

Excerpt 4.1.6

1 Int: *a hedgehog*

2 Clinton: ***hotcha***

3 Sp2: *they all reckon they eat hedgehogs you know*

4 *but they don't know that's not 4the case*

5 Clinton: *it is*

6 Spk2: *you don't do you*

7 Clinton: [I] *have had (eaten) hundreds*

8 Clinton: *yeah and Gypsy people would never tell I'll tell now*

9 *but a Gypsy would nevertell gordgers how to do 'em*

10 Clinton: *they used to say to pluck 'em out and then do them in clay*

11 *that's a load of bollocks load of bollocks I could do an hedgehog in no time and*
12 *give you some of it and you'd think fucking hell what's this is it chicken or what*
13 *its better than chicken though*

Within excerpt 4.1.6 Clinton also related another repeated theme, that Gypsy culture was commonly not shared or that within the GRT communities many would not wish to share cultural information (Okely, 83; Matras, 2010). In this excerpt, Clinton considered this information as something '*Gypsy people would never tell I'll tell'*. This epistemological proposition revealed Clinton as providing knowledge from someone who had authority to do this. This reinforced the interactants within the exchange and positioned the information as genuine, although it was unclear if this was not with humorous intention as there was an effect of hyperbole suggested in the profanity and exclamatory use of '*that's a load of bollocks load of bollocks'*. Okely (1983) argued a distinction within Romani culture symbolism between outside and inside. These ideas symbolising an ideological difference that distinguished lines between physical space (outside and inside the home) as well as mental, ethnic, and cultural lines between those within the ethnic group and those outside the community. In this sense, the cultural values could be seen to lie within the ethnic group and were kept pure and inviolate by not sharing with those from outside the community. However, in this interaction, there was a suggestion that Clinton, in his circumnavigation of this label, was, actually, mocking this idea with the playful intention of challenging this particular stereotype as well as other associations. This playfulness suggested a confidence and reinforcement of his identifying himself as an Anglo-Roma.

The variants for 'hedgehog' are found across several archival sources with the exception of Ben's *tikni bawlo*. This looks to be an innovation as *bawlo* has been recorded as 'pig' in European Romani (Manchester Project, 2021) and *tikni* (*tickno*) as 'small'. This was not found in the Anglo-Romani word lists. Those produced by Clinton, Duncan and Jake corresponded with those words listed within the Anglo-Romani source lists, however. The variable 'fire' was consistent across participants as *yok,* whilst Ben produced both *yog* /yɒg/ and *yag* /yæg/ The phonological distinction is found across several sources and represented an alternative realisation. Multiple variants for a single variable indicated exposure to more than one dialect. Given Ben's background (section 3.1.1), it would be expected to affect the degree to which this

would influence his response. Finally, in terms of comparable variants, 'Rain' defined as *brushindo* (noun) and *bresheniskri* for 'rain' and 'rainy' respectively from participants Jake and Ben. Again, Jake's production of *brushindo* was found across archive sources and corresponded with *balval*, the European Romani word (Manchester Project, 2021). *Bresheniskri* was not found amongst Anglo-Romani vocabulary lists and indicated a degree of language proficiency in terms of the morphology represented by this variant's production by Ben.

The remaining variants were from Jake, who produced variants for 'hay' *Kas*, 'sun' *Kam*, 'wind' *Baval* and 'earth' *Ful* within the subcategory, Nature and Weather. Mentioned previously, Jake demonstrated a pattern of elicitations that could be interpreted as relating to agriculture and farming, corresponding with the cultural contexts of Lincolnshire, being the county of his residency. This would need further investigation and was a limitation of the current study.

4.2.4 Customs

The final subcategorization within the SRN The Outside World 1 was 'Customs'. This was the smallest subcategory in terms of elicited variants. This was, perhaps, unsurprising as this was a relatively obscure and opaque semantic area with the least definable parameters relative to areas such as 'food' or 'nature'.

Table 4.1.1 SRN:The Outside World 1: Customs

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Traveller/Gypsy	*Pavee Lakin*		*Traveller/ Gypsy*	*Rumnus gearer/Rumnus* MC, DV1, IV, DV1, DV3	*Romani* T1, MC, S1, W1, SC1, (MC, T1, L1)	*Rom*
Gypsy Fair					*Velgooras* R, T1, W1	
Traditional Wagon				*Vardo*	*Vardo*	*Vardo/vurdon*

				DV2, DV1, DV3	DV2, DV1, DV3	DV2, DV1, DV3/S3

Table 4.1.2 Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

Jake produced three variants from three of the variables, with Ben being less productive within this subcategory with only two variants. As a small subcategory there would be expected to be far less variation, and, therefore, fewer tangible differences in terms of productivity, which would be accountable as a factor of identity given their backgrounds as outlined in section 4.1.1.

In excerpt 4.1.7, Ben discussed cultural practices and beliefs of his community, which he self-identified as *Rom*. Transcription conventions (Appendix 8) did not apply to the following written response

Excerpt 4.1.7

1 Ben: *One that all **Rom** used to stick to was don't tell outsiders (non-**Rom**)*

2 *any personal things.*

My father was strict about this, he also was very strict with my younger sisters as in boyfriends and the like

I was allowed much more freedom than my sisters. I didn't do well at school and wagged much of it, only my mum was a bit concerned. but with my sisters they did really well at school.

Also, food could be an issue with my father. He would carefully check meat and how it was cooked. He would not allow anyone to use his cup etc.

Total strangers were not made welcome unless with a relative.

In this excerpt taken from the email interview with Ben, the term *Rom* was used to describe his cultural background, as an endonym, referring to his cultural or ethnic identity. In this example, Ben narrated an account of growing up as a Roma with reference to his experience of being raised by his Romani father. He indicated modality with his commitment to positionality with 'all Rom' and continues to outline a number of propositions within this propositional stance such as '*don't tell outsiders'* with his father being '*strict about this' and* '*very strict with my younger sisters as in boyfriends and the like.* Within the same narrative he also stated his own degree of affinity by declaring his own, and potentially others' stance with '*all **Rom** used to stick to [this]'*, indicating an evaluation that not all or many did not continue to follow this characteristic. He continued to describe his own identity with the actions of his father, including certain culturally associated actions. The distinction outlined by Ben in his narrative between the position of women indicated Ben's evaluative stance towards the distinctions made in his narrative between men and women in what he outlined as his experience of Romani culture as he stated: '*I was allowed much more freedom than my sisters'*. The cultural proposition of the distinctions made for men and women in Roma societies was also referenced in a number of sources showing that men and women were ascribed distinct roles, which was one that Ben conformed to (Okely, 1983). Simlarly, the notion of cleanliness was also mentioned in this excerpt in '*He would carefully check meat and how it was cooked'* with the concept of the distinction between 'inside' and outside' in '*He would not allow anyone to use his cup etc.'* as well as *' Total strangers were not made welcome unless with a relative.* The propositions in terms of content and selection indicated the characteristics that Ben considered contrastive to those ideas as understood by his audience (the interviewer), and his cultural norms and values

currently, and represented a retrospective view of what was acceptable within his Romani speaking family.

Self-referent terms varied between participants, as can be understood with Duncan's use of *Travelling/Gypsy man*, and *Traveller* as a self-referent terms, whilst the term *Gypsy* for Ben was used to collocate with 'language', 'community', 'tribe', 'people', 'race', 'warriors', 'slaves', 'wagons', 'smith', and 'breeds' established in the interview data. The term *Gypsy Traveller people* was used by Clinton within general conversation, but when asked directly for a self-referent term, the variant *Rumnus gearer* was evoked. The variant *Rumnus* without the collocated *gearer* (Anglo-Romani for man (Dawson, 2002)) was found in the archive source, as is *Romani*, drawn from Jake. The significance of the self-referent was framed in the degree of variation between community groups and individuals' use of the endonymic terms. There was a degree of controversy regarding the terms used to refer to one's own community or one's own cultural identity, with this being, markedly, the case for application of the terms 'Irish Traveller' or 'Anglo-Romani' (see chapter 1 section 2.1.2 for a detailed discussion of the literature). The use of the ethnic referent was, also, referred to in section 5.3 in 'Phrases'. On several occasions Duncan alternated between 'Gypsy man', 'Travelling man' and 'Romani' to refer to his own ethnic identity, and illustrated flexibility existing with the self-referencing of his own ethnicity. For the purpose of this study, as mentioned in section 3.3.1, the participants self-ascribed identity, thereby, producing an emic, self-descriptive study.

The variants for 'wagon' were consistent between participants and realised as variant *vardo,* whilst Ben produced the additional variant *vurdon*. Both variants were found in the archival data, whilst *vurdon* was from a much earlier source, and, therefore, an isolated single source entry. Jake produced the variant *velgoora* for fair or Gypsy fair. This was, also, extracted for a community meeting as part of the SRN Outside World 2. This illustrated an example where dialect variants occupied multiple contexts of use and another example of extension (see also excerpt 4.2 and Table 4.4 in the case of Jake's use of *livna* for beer).

4.3 Conclusion

Within this chapter concerning lexical variation, factors that influenced and affected individual variation, with associated acts of identity within interactional dialogue have been examined and discussed. In terms of lexical variation there were a number of salient areas. For distinction between self-identification of ethnic background between Irish and Anglo-Romani, there existed a complex relationship between lexical variation and background. Participants' lexical variation was found to be associated and distinguished by that ethnic association. Self-referent

cultural associations were established as determinants of lexical production and usage. In the case of Judy, there was a nuanced distinction between lexis historically associated with Irish Traveller communities and non-standard forms of English, as Judy employed non-standard forms of English mixed with Cant lexis as part of her dialectal repertoire. In the case of Duncan there was a distinction between this Irish Pavee Cant lexis, and that associated historically, and as self-reportedly used within the Anglo-Romani community of practice. Variants used were often terms that referred to a number of concepts, such as, *oben* referring to 'food' as well as vegetables, and this was observed as a general pattern that might be associated with minority language use. In relation to this, a number of functional words were also detailed as being retained although their complexity was observed as reduced in terms of the number of terms used between participants, and again there was variation in relation to background factors relative to exposure, and current use within communities. Diatopic factors or locale was suggested as being an influencing effect across production of a number of variants and this was again referred to later within the analysis chapters. A number of examples relating to phonological features, as well as lexical characteristics, relating to meanings of words associated with regional traits were also considered in relation to this as a characteristic amongst speakers. Words, such as, that for 'potato' were considered in terms of their cultural and regional saliency. Individual variation relevant to productivity, innovation and inflectional use have been, also, found in asscociation with background factors. In the case of Ben and Jake, examples where inflectional endings and comparability to European Romani words were greater than with other participants, with background factors being considered influential for this variation. Observation between archival sources has revealed a degree of comparability and enabled and enhanced those observations, showing structured patterns of variation between the participants of this study.

In terms of interactional data, identity as a sociocultural phenomenon has been highlighted with examples. Indexicality of identity categories was demonstrated through a number of cases. Cultural phenomena, such as, reference to horse culture and the relevance of this within interactional discourse has been observed and discussed in terms of its relevance to discussion of lexical variations as acts of identity. Emergent insights into acts of identity have been noted through for example of the use of pronouns within interactions, and how this created notions of Gypsy, Roman and Traveller identity through interactions within the interview excerpts. The use of self-referent terms has also been discussed, marking a degree to which identity was not a fixed state, and the extent to which such notions are transitional within the dialogues of

the interviews analysed. The next chapter examines a second semantic category in consideration of those influencing factors for lexical variation within the semantic field of ‘The outside World 2’.

Chapter 5 Analysis: Outside World 2

5.1 Semantic category: The Outside World 2

The SRN 'The Outside World 2' has been divided into four relational categories. These are: 'Jobs', 'Places', 'Money', and 'Numbers and Time'. See chapter 3 section 3.5 for an outline of the development and selection process for the semantic categories for this SRNs. From fifty variables ('standard English' prompts) a total of 81 variants were elicited across all informants for this category. This SRN was the second most productive for each of the informants. What follows is an analysis of variation in order of the most prominent variations to the least. As previously stated, the analysis has focused on lexical variations, and, where relevant, phonological variations. Those variations will be discussed in relation to background and contextual factors in order to compare those variants with archival materials accessed and collated as part of this project as outlined in chapter 3, section 3.5.

5.1.1 Jobs

The semantic subcategory within 'Outside World 2' that drew the highest frequency of variation was 'Jobs'. Vocation-related variables 'doctor', 'construction worker', and 'police' gave the highest degree of variation in response to corresponding standard English variables. Table 5.1 has depicted responses to standard English variables related to 'Jobs'. As mentioned, additional words not only SRNs produced by participants were also added. See chapter 3, section 3.5 for a description of the selection process and lexical categories. In summary, these were based on a selection of words from literary sources of Anglo-Romani origins (e.g., Manchester Project, 2021).

Table 5.1 SRN: The Outside World 2: Jobs

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Housewife				*Ker Monishina* DV2	*Monashin*	*Keri Mort*
Police	*Shades* DV4		*Mingerer/ Gavver/ Muskerer* /DV2/MC ,DV2	*Mushkerers*	*Muskeroes* FS	*Mushkro* FS

Doctor				*Travingerer* MC	*Bara /Bawra/ Drabengro* /_/DV2, D1, DV1, S1, B2, SC1, MC	*Drabeskro/ drabmenagri*
Soldier				*Kuddyman*	*Koolie* B2	*Yogkosh engro*
Manager	*Capta in/ Gaff*			*raior*	*Sherengro* S1,W1,B2, SC1	*Baro sherro mush*
Construction worker				*Buttier*	*Ker Mush*	*Keravano Mush*
Worker				*Buttier*		*Booteskro*

Table 5.2 Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

Table 5.1 revealed variable 'police' elicited the most variants. Evidence was the phonological variations between these elicited items. Although these variants possessed similar characteristics in terms of phonology, (*Muskerer /muskəɹə/* and *Mushkro /muʃkɹau/*) variants *gavver* and *shades* weredistinct in terms of their etymology. The word *shades* was distinct in terms of dialect and origin and produced by Judy. Judy described herself as having an Irish Traveller heritage and background (see section 3.3.1). The dialectal variant *shades* was listed within archive sources, that were collected from those who could be described as Irish Pavee Cant speakers. This term was recorded by Dawson (2011) as Irish Traveller Pavee Cant (see section 3.8.2 for discussion on use of folk-linguistic sources) as attested by self-characterized speakers of Traveller Cant. Dawson's (2011) list collated lexis from primary and secondary sources from Dawson's own recordings and older sources (Leland and MacAlister, 1937). Duncan produced three separate words for 'police'. Duncan discusses his heritage as being both Irish Traveller and Anglo-Romani (see Table 4.3) and this he explained, was a reason for his understanding of words associated with Irish Traveller and Roma. The transcript below in excerpt 5.0 illustrated Duncan's explanation for the variation that existed between the variants *mingerer* and *gavver*.

Excerpt 5.0

1	Duncan	*there's two words for these… one of them the oldest words I would say* ***mingerer***
2		*or I would say there's another word I would say*
3		*which is I'm just trying to think of it now off the top of my head*
4		*it's uhm a* ***gavver*** *and the reason I think* ***gavver*** *is to* ***gavver*** *something do you*
5		*understand me*

Duncan continued to offer a 'folk-linguistic' account of the etymology of the word *gavver*. The analysis of folk-linguistics regarded non-linguists' and speakers' understanding of communicative codes as a means for understanding knowledge, attitude, and ideological beliefs (Niedzielski and Preston, 2001). The approach has been debunked by some as potentially inaccurate or misleading (Kroskrity, 2009) as words might be inaccurately attributed to an etymological root, with notions of nationalism or ideology dominating. Whilst others noted the associated strengths of social meaning behind metalanguage discussion (Albury, 2017; Rieder, 2018), giving voice, empowerment and agency for speakers. The excerpt below illustrated how Duncan has provided details of lexical etymology. As in excerpt 5.0, Duncan offered a number

of folk-linguistic accounts within his discussion of dialect and heritage, placing his propositional knowledge as having authority within the discussion. Folk-linguistic accounts, whilst not always corroborated, offered ideological perceptions of Romani dialect, which indicated more about life experience and situation of each participant than 'expert' views of linguists alone (Rieder, 2018). Duncan provided a tentative account of the etymology of the variant *gavver* shown with the use of modality 'I think' in '*and the reason I think gavver is to gavver something*'. Duncan's narration provided detail about the origin of the word *gavver* within the interview, and, by doing so, revealed himself as knowledgeable. This understanding is authenticated by the implication of his self-describing as Irish Traveller with Roma background or appellation. The notion of authenticity is co-constructed within the interview as Duncan was tasked with providing personal understanding and knowledge of dialect and lexis. Duncan hedged his propositions with reference to others' 'expert' opinions as well as through use of 'I think' to show his stance through use of this modality.

In excerpt 5.1, Duncan continued to explain the word *gavver.*

Excerpt 5.1

1	Duncan	*so they [the police] would **gavver** you… so **gavver** I think that's where it came*
2		*from I don't know but that's one word… **mingerer** or **muskerer**… there's lots of*
3		*words for that one word so **gavver***

In this excerpt, Duncan continued to offer a meta-linguistic account of the origins of the word *gavver* and produced the variants *mingerer* and *muskerer*, whereas Duncan explained the origin of these dialectal variants as linked to both Irish Traveller as well as to Anglo-Romani speakers' identities This arguably acted to underline his ethnic heritage as a factor for his knowledge, use and understanding of the lexical variants.

The plural form *gavvas* has been recorded in the Manchester Corpus (2021) and *gavver* was a singular noun variant recorded in the current project's data. The item has been recorded as European Romani, which was derivation of the verb *garav* meaning 'hide' (Manchester Corpus, 2021). There were several variants for 'police' within the Anglo-Romani corpus and in the archive sources collated for this project. One other record of the singular noun variant *gavver* can be found in Dawson's Derbyshire list (Dawson, 2002). The morphological plural marker 's' was more common as 'police' was used in its plural form. The variant demonstrated the morphological integration of English plural suffix into Anglo-Romani lexis (Matras, 2010). Duncan also produced *mingerer*, similar to *mengra,* recorded in the Derbyshire lists of Dawson

(2002). This variant has been found in the records of Welsh Romani, and was an abbreviation of *prastamangra* (Sampson, 1923). The variant *mingerer* used by Duncan, did not appear in the archive or corpus. Assimilation might explain the variants *mushkerer* and *mengra* as the ending *-gra* superseded the ending *-rer* and intuitively these variants were more likely to be linked etymologically than to *gavvar*. Examples of assimilation patterns can be found relating to language contact across the dialects of Romani and specifically related to morphology and phonological changes (Benisek, 2020 : 27).

Four other variants were elicited from Duncan, Clinton, Jake and Ben for 'Police'. Phonological and morphological distinctions were evident. The initial postvocalic consonant was distinct between variants. There was an alveolar fricative /s/ in *muskerer /muskəɹə/* and post alveolar fricative /ʃ/ in the case of Clinton and Ben in *mushkerer* /muʃkəɹə/ and *mushkro /muʃkɹəʊ/*. The realisation *mooshkero* was recorded by Crofton and Smart (1875) from two informants in Nottinghamshire identified as Weston Boswell and Isaac Heron. *Mushriika* was recorded in the sources found in the Leeds Special Collection (1979) (a posterity recording of both Pavee Cant and Anglo-Romani speakers). The source was an archival recording of both Cant and Romani speakers and based on interviews recorded for archive purposes for the maintenance of Romani and Cant dialect. There were only two variants recorded that were realised with a post alveolar fricative from archival and corpus sources. This represented a less common feature, and, therefore, hypothesized as a regionally influenced phonological variation notwithstanding limitations (see section 3.8.2). This has been suggested to relate to diatopical or locational characteristics of individual variation and a focus of the analysis.

The suffix *-er* in *muske**r**(s)* was used by Duncan and Clinton, whilst the suffix *-ro(es)* was used by Jake and Ben in *muskr**oes*** and *mushk**ro*** respectively. This English suffix '-s' was used as an indication of plurality. The plural 's' ending is found within the Derbyshire list recorded as *muskaeros* and recorded as 'old Welsh Romani' (Sampson, 1926a) with the example recorded as *mushkros*. The variant *muskerer(s)* recorded for Duncan and Clinton was not found within the archive or corpus and again might be represented as a regional variant. The following except from the interview with Clinton marked an awareness of the differentiation between Romani lexis and regional vernacular. The participant's background was described in section 3.3.1 with each participant's heritage profiled within each interview. As mentioned Speaker 2 was an acquaintance of Clinton, and had not discussed the topic of ethnicity previously. In excerpt 5.2 below, Clinton tentatively stated the term *mushkeras* had a dialect origin related to

ethnic status, whilst the interviewer contradicted that and claimed its status as 'slang'. The following exchange related to attitudes towards the status of words and their ethnic origins.

Excerpt 5.2

1 Clinton: *police* ***mushkeras****…*

2 Interviewer: *that's one that we grew up with*

3 Clinton: *yeah but I think that's one of the words that come of us*

4 Interviewer: *yeah been borrowed*

5 Speaker 2: *Town Town slang innit*

In this exchange, Clinton responded to the claim of the interviewer that '*that's one that we grew up with'* by disputing '*that's one of the words that come of us'* hedged with '*yeah but I think'*. In the interaction, Clinton was reaffirming the variant '*mushkeras…*' as originating and '*coming from..*' a Roma origin. Clinton also reaffirmed his position of ethnicity by using the plural pronoun '*us*', reflecting an ingroup stance in proposing his ethnic status as distinct from others within the conversation.

The variable 'doctor' was extracted from the six variants. Variation was evident between Clinton, Jake, and Ben. Both Clinton and Jake were from the East Midlands within the UK. As mentioned in section 3.3.1, Ben has spent a significant amount of time in New Zealand and provided a contrastive set of data in relation to his background. Ben identified as Roma with Welsh ancestry, with some but limited association with the East Midlands. Most similar were variants *drabengro* and *drabeskro*. The suffix *-gro* and *-kro* have been recorded as an early inherited Romani (Matras, 2002: 74), and intensified general meaning of relation or pertinence (Koptjevskaja-Tamm, 2000; Meli, 2016). The Romani genitive suffix *-kero* and allomorph *-gero* have been added to a root stem as a strategy for new lexicon production (Scala, 2020). In terms of morphology, the general patterns connected a stem ending with a masculine plural stem *-n* taking the allomorph *-gero*. Both those endings appear in the project data for the variable 'Doctor', whilst within the online corpus and archival sources *drabeskro* did not appear. However, *Drabengroe* did appear in the Derbyshire word list (Dawson, 2002) and appeared in five other sources within the archival material indicated in Table 5.1. It was also recorded within the online corpus (Manchester Corpus, 2021) as *drabbengro*. *Drab* was

European Romani for medicine and *drabengro* and variant *drabeskro* were derivatives of this root (Manchester Corpus, 2021).

Clinton produced the variant *travingerer /tɹaviŋgəɹə/*. This was a phonological variation of the variant *dravvingra /dɹaviŋgɹa/* recorded in the Manchester corpus and *drabengroe* /dɹabeŋgɹəʊ/ listed in the Derbyshire word list. *Bara* and *bawra* were obtained from Jake for 'Doctor'. *Bawra* was listed in the Derbyshire word list (Dawson, 2002) as meaning specialist doctor although the variant 'bara' was not listed. The word *baro* meant 'big' in European Romani (Manchester corpus, 2021) and was used to form a compound in the case of *bawra drabengro* meaning surgeon or specialist consultant. Based on the corpus, archive, and interview data this represented a low frequency term and indicated a higher degree of proficiency and knowledge of lexis evident in the degree of inflectional morphology and productivity further outlined in chapters 5-7 and characteristics of Jake and Ben's repertoire (see Table 4.0).

'Soldier' produced variants for informants Clinton, Jake, and Ben. Clinton's variant *kuddyman* was not recorded in the archive sources or Manchester corpus, and represented a unique variant. Excerpt 5.3. below was a conversation where Clinton contextualised the use of *kuddyman*.

Excerpt 5.3

1	Clinton:	*there are alot of jokes like **dick** to the **kunny gearer akai** that means*
2		*look at the army man or airforce man a **kuddy** means army or…*
3	Interviewer:	*a **kuddy** man*
4	Clinton:	*it means army man or airforce man you know*
5	Interviewer:	*what someone in uniform?*
6	Clinton:	*when they used to have a lot like during the war and that and*
7		*they were stationed in different places and they was all over the place*
8		*then they'd say **dick** to the **kuddy***
9		*or he's up to something or trying to do something you know right*
10		*I'll tell you story*

In this excerpt, Clinton initiated the narrative by explaining the use of the word for 'soldier'. The example phrase '*dick to the kuddy…'* or 'look at the soldier' suggested an affective stance towards the actions of soldiers as in the following '*he's up to something or trying to do something'* implying some kind of bad intention. This reference contextualised a playful intention to introduce a humorous anecdote in order to illustrate the use and meaning of the word '*kuddy*'. Clinton's attempt to convey the meaning by way of a humorous anecdote with hyperbolic modality '*they was all over the place..'* with a dramatic introduction in: '*he's up to something or trying to do something you know right I'll tell you story…'* created intrigue and tension to foreground the contextualisation that followed.

Koolie was produced by Jake and *kuddy /kudi/* and *koolie /ku::li/* were phonological and semantic cognates. *Koolie* can be found within the archival data recorded by Borrow (1874: in Robert Dawson's Special Collection) and might represent a diminutive version of *kooromengro* attributed to Crofton and Smart (1875). *Kurramangra* is referenced in the Manchester Corpus, and cited by Russel (1916). The European Romani word for beat is *kur* (Manchester Corpus, 2021) and was an attributable root for these variants. The English compound 'man' and morphological suffix *-ie* in *kuddieman* and *koolie* were evident in the formation of the variants Clinton and Jake both produced. Ben responded with variant *yogkosh engro* that did not appear in the archival sources. As mentioned, for European Romani the genitival suffix *-engro* was affixed in cases where the stem ended with *-n* for masculine plurals. The word *yogkosh engro* literally translated as a person of fire stick/wood. This variant revealed the divergence between variations for Ben and the other informants within the study. This variant might be considered relevant to Ben's Welsh heritage as well as his time in New Zealand and could be an example of lexical innovation.

'Construction worker' and metonym 'worker' produced a number of responses for Judy, for Graham, and Duncan. *Buttier* was not found within the corpus and archival data although *butti* has been recorded by Griffiths and Yates (1934) for 'English-Romani'. *Butiyokkeris* was a close cognate recorded by Winstedt (1948) with the stem *buti* meaning 'work' in European Romani (Manchester Project, 2021). As discussed in 3.8.2, there was variation in interview contexts. This might have had an effect on performance pressure or factors related to attrition, which might have hampered recall, whist many variants as in the case of *buttier* below in excerpt 5.4 the response was given without hesitation.

Excerpt 5.4

1 Clinton a ***buttier*** that means worker

The difference in variety spoken by Ben was again apparent in the production of variant *booteskro* for worker. This variant can be classed as cognate (aligning with European Romani Manchester, 2021). The stem *buti* was used for work. However, the suffix *-er* for Clinton followed the English derivational suffix *-er* and used as a productive strategy. Jake used the Romani suffix *-kero*. Jake and Ben used the European Romani variant *ker* ,which can be interpreted as 'to make' or 'build' for 'construction worker'. *Ker mush* produced by Jake and *keravano mush* for Ben did not appear in the archive or online corpus. This could be owing to limitations of sources used for this study or these might represent variants that were, as yet, unrecorded or lexical innovations.

Following a similar pattern, Judy, Graham, and Duncan did not produce responses for a variable for 'housewife', presumably, because it had dropped out of use. This was intuitively a high frequency lexical item, although not produced, that characterised the productivity of Judy, Graham and Duncan as distinct from Clinton, Jake and Ben. Within that SRN, the term for 'housewife' was categorised in relation to vocation. Similarly, words for 'Wife' in the SRN category 'Ages and Relationships' were also limited. This might indicate the low frequency of use more generally for informants for those terms within those varieties. It was also possible words that were gender specific terms were maintained in differing degrees which would need further investigation.

In excerpt 5.5 below, there was also a degree of hesitancy during the conversation discussing the variable 'wife' which appeared related to the angling of the question with respect to the use of the gender neutral term 'partner'.

Excerpt 5.5

1 Interviewer: *partner as in husband and wife…*

2 Clinton: *no I don't think there's any words for that…*

3 Interviewer: *…yeah*

4 Clinton: *for my wife this is my* ***monishna*** *and that means my wife*

That feature of Anglo-Romani demonstrated that, at least, for Clinton's variety gender neutral terms were not attributable or been adopted or previously featured (Matras and Tenser, 2020).

Within the archives comparable lexis have been recorded. The Derbyshire list (Dawson, 2002) recorded *monishna*. The word *monishee* was recorded by Way (20th Century). Those terms were examples, amongst other variants, that have shown persistence across several generations of English Romani speakers. Jake produced the comparable variant *monashin* that possessed a clear cognate form *manushi* (Borrow, 1874: in Robert Dawson's Special Collection). The variant elicited by Jake *monashin* was not recorded in the archive or online corpus although was a cognate of the eleven variants found recorded amongst archival sources. This could be a variation associated with the regional area of Lincolnshire or an individual/group variant. Ben produced the variant *keri mort*. This did not appear elsewhere recorded as Anglo-Romani. This might be derived from a Welsh variant in relation to the ancestry of Ben. Dawson (2009) listed the term *kairikeni* a Scots/Irish/Romany Traveller term meaning housekeeper which might be a related term. That revealed a degree of cross-over in relation to lexical variation and ethnic identity ascribed to the speakers of those taking in part in the data collection of the archival materials.

For 'manager', Jake and Ben produced *sherengro* and *baro sherro mush* respectively. The former produced by Jake was a Romani word and appeared in sources both in the archive material and online corpus. The Anglo-Romani variant term *sherikeno gaero* appeared in the online corpus cited as No author (1929). The cognate *šero* also appeared in the Manchester corpus (2021) and was defined as being of European Romani origin meaning 'head'. The term used by Jake used the suffix *-engro*. That variant was found in four archival sources (Sanderson, 19thC; Way, 19thC; Borrow, 1974: in Robert Dawson's Special Collection; Crofton & Smart, 1875). The remaining variables within the category 'Jobs' produced fewer variants across the participants. Ben used a lexical innovation and compound *baro sherro mush or* 'big head man'. It has been speculated that Ben elicited masculine forms since *-o* was used to indicate male referents in languages, such as, Spanish. That term did not appear in the archive or corpus, and might represent an idiolectal form or dialect variant not represented in the corpus of the archive. Clinton produced *raior* and has been exhibited below in excerpt 5.6 below.

Excerpt 5.6

1 Clinton: ***raior*** *really you heard that one…*

2 *… **raior** means you're the boss…*

3 *…yeah if I come on a farm and you two was on a farm and I wanted to see*

4 *the man who was in charge or something who's going to sell me the scrap*

5 *or whatever **dick** to the **raior** that means look see the boss…*

The use of the variant, in this instance, in excerpt 5.6 indicated the term *raior* referred to *the boss….* . Within this narrative Clinton referenced the vocation 'scrap' or selling of recyclable metal. The use of Anglo-Romani lexis demonstrated an affective stance towards hierarchy and an attitudinal stance towards professional status. The term /ræjɔ/ is found within archival sources but not used for 'manager' or 'leader' and has been presented as a unique elicitation within this project's data. The term *rai* (raj) is found in European Romani, and defined as 'gentleman' or 'lord' (Manchester corpus, 2021), *raia* (Sampson, 1911))

Judy was less productive overall in terms of non-standard variants that she identified as Irish Pavee Cant related dialect. Judy expressed that discussing Irish Traveller culture was problematic in terms of disapproval from other community members. She produced a total of 38 nonstandard terms, a large number of which could be found within previous archival sources as related to an Irish Traveller ethnolect. The term for 'manager' elicited *captain* and *gaff* although those were not found within the literature associated with Irish Traveller dialect. Those terms have an etymological root in 14th and 15th century French terms, respectively (Oxford Dictionary, 2020). Notwithstanding, those terms represented nonstandard variants used by Judy. This suggested dialect might be a less prominent feature in the case of Judy, whilst Judy's variety was distinct in terms of her own unique self-described identity in her use of lexical variants. This might also illustrate how nonstandard forms have merged to a point where Irish Traveller dialect and other nonstandard lexis, more generally associated with English vernacular variation, have merged. This was something discussed by Dawson (2011) in relation to Romani and Irish Traveller Cant, and was evident here for Judy in relation to English nonstandard variation. The ethnic Identity Score Index was low at four out of a possible 14 (see Table 4.2). The regional identity score was also low at four, and, possibly, a factor in relation to productivity. Visual and Attitudinal scores were not collected following ethical considerations (see section 3.8.1 for ethical considerations). Those low scores would suggest a more modest production of nonstandard forms. Background factors were discussed in relation to other contextual considerations and in relation to factors relative to other informant scores (Underwood, 1988; Llamas, 1999).

A number of factors might influence productivity, and, therefore, maintenance of particular items. In terms of generational differences, Clinton, Jake and Ben were from an older

generation which might also be an influential factor. This was explicitly discussed by Clinton and reported by Ben. Both were born mid-20th century or 1950s and reported being raised with Romani as a family language. Judy, Graham, and Duncan were two decades apart in terms of age, born in the 1970s, which might be a significant factor in determining the number of lexical variants reported by each informant (see section 3.3.1).

5.1.2 Places

The second most productive category within the SRN 'Outside World 1' was 'place. There were a total of 25 variants elicited for English variables associated semantically with 'place'. This proved a more productive field and indicative of higher frequency of usage, and maintenance for this semantic category. Table 5.3 illustrated the items elicited for the lexical variable 'house', 'fire (place)' and 'here' being the most productive variables, whilst the variables 'caravan' and 'main room' were the least productive in terms of variation. Those were all chosen based on the selection procedure outlined in section 3.5. The variables 'long journey', 'toilet', and 'Lincolnshire' were not items in the SRN, and were elicited as 'other' through unrehearsed conversation. 'House' was notable amongst the list of variables as a variable that elicited variation between all informants. The term 'here' was conspicuous as a deictic term and categorised, in this instance, in relation to its conceptual association with 'place' albeit with its referent point specific to the context of its use in relation to the speakers.

Table 5.3 SRN: The Outside World 2: Places

Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Road			*Drum*		*Drom/Drum* MC/	
town				*Gaff*	*Leve/Gav* / DV2, N1, R2, F1, W1, B2, SC1	*Gav* DV2, N1, R2, F1, W1, B2, SC1
Long Journey					*Lango Drom*	
House/home	*Tints*	*Kenner* DV2,MC	*Atchin Tan* MC	*Ker* DV1,T1, N1,F1, S1, W1,	*Atchin Tan, Tan* MC/MC	*Ker* DV1,T1, N1,F1, S1, W1,

				B2,SC1, L1. LSP		B2,SC1, L1. LSP
Caravan		*Trailor*		*vardo*		
Toilet		*Muttering Kenner*				
Main room of house, At back of Wagon				*Ker*		
Here				*Ochiakai*	*Aki*	*Akai* D1, S1, GY, G
Outside					*Aurial*	
Lincolnshire					*Rushni Tem*	

Table 5.4 Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

Table 5.3 revealed two pairs of participants used the same variants for 'house'. Duncan and Jake used the term *atchin tan*, whilst *ker* was also recorded for informants Clinton and Ben.

The variant *tint* was used by Judy and was distinct as not having an associated match within the archive. Graham used the word *kenner* as a compound and contextualised its use, in defining t *muttering kenner* as 'piss house' shown below in excerpt 5.7. Graham was less productive in terms of numbers of variants produced as shown in Table 4.1 in chapter 4. Graham did not complete all parts of the SuRE interview as outlined in Table 3.0 and discussed in section 3.8.2 in terms of limitations. Graham was previously described as ascribing to a Roma identity, which was established and discussed in terms of his identity and production of phrases and variants within interactions.

Excerpt 5.7

1 Graham: *so you could say if I were in a pub or I'd say I'm* ***geling*** *to the* ***muttering kenner***

2 *or* ***parnying*** *the* ***grai*** *it means I'm going to the piss house*

In excerpt 5.7, Graham explained the use of an idiomatic metaphor in '*parnying the grai'*. This was indicative of a lexis embedded within English as a matrix language (Muyskin and Milroy, 1995), with the significance of the use metaphorical or innovative use of language to express meaning. Graham used the personal pronoun in the phrase '*if I were in a pub or I'd say'* showing an affective stance towards the use of a Romani phrase as part of his own use of dialect. Graham exhibited a degree of nuance in relation to his position as a speaker of Romani dialect as he positioned himself as using Anglo-Romani explaining how he would use the words discussed.

Below in except 5.8 Duncan explained the different use of *atchin tan*, giving two different contexts for the use of the word.

Excerpt 5.8

1 Duncan: ***tan*** *is a house or it could be a you know a building or a piece of land*

2 *but it would be to* ***atch*** *…* ***atchin tan*** *… a stopping place*

Again in this excerpt (5.8.) Duncan authenticated his position with his explanation of the term tan with '*tan is a house'* demonstrating an epistemological stance towards this assertion, followed by the extension '*but it would be to* ***atch*** *…* ***atchin tan*** *… a stopping place*' . The meta-linguistic explanation placed the interviewee as an informer, and, also, affirmed his status within the interaction, in addition to further illustrating an affective stance towards the importance attached to the use and meaning of this lexis within that context.

Table 5.9, also, mentioned the term *ker* used by Clinton for 'house'. During the conversation with another speaker present, speaker 2, during the interview, questioned the use of *ker* for house and whether *ker* would be used for caravan. The distinction between the use of *ker* and *vardo* was emphasised by Clinton.

Excerpt 5.9

1 Friend: *that would have been a caravan in its day… would you call a caravan a* ***kear*** *then?*

2 Clinton: *its where I live… you call where you live a* ***ker****… it's an house isn't it…*

3 *I mean then it would have been a* ***vardo****… a* ***vardo*** *would be a caravan*

4 *or wagon yeah…*

In excerpt 5.9 Clinton responded to a friend who suggested living in a caravan the type of dwelling Anglo-Romani or Travellers would reside in historically by the reference '*in its day'* referencing '*would you call a caravan a* ***kear*** *then*'. Clinton replied by indicating *vardo* was the term for caravan. Clinton's response indicated he was challenged by the assumption that a 'house' would equate with 'caravan' in his reply '*it's where I live… you call where you live a ker… it's an house isn't it…*'. Clinton's response revealed an evaluation of the friend's stereotype as he corrected their understanding of the use of the word *ker* as '*ker*' and '*vardo*' as having distinct meanings. Clinton challenges that assumption with the use of the tag question '*it's an house isn't it*' emphasising the recognisable degree of distinction between 'house' and 'caravan'. That also acted to challenge the underlying assumption and stance towards the potential threat of a subsidiary narrative proposed by the friend.

Within the online corpus and archive material, *tan*, *ker* and *kenner* were found recorded for 'house' with the example '*atch something in the tan*' (Manchester Corpus, 2021). Duncan also used *atchin tan* for 'camp' or 'stop over'. The meaning of *tan* was recorded as 'place' and *atch* as 'stop' in European Romani with the term meaning stopping place (Manchester Corpus, 2021). The use of *atchin tan* for 'house' was not found in the archival material although used for house within the records and data base in the Manchester Corpus (2021). Clinton and Ben used the variant *ker*. Both Clinton and Ben, as mentioned, were generally more productive although generational differences (Table 4.1) were a factor. Jake, also from an earlier generation, also used *atchin tan* but not *ker*. Age did not appear to influence variation for certain variants, at least observable in the data, although owing to limitations (section 3.8.2) more data

would be necessary to make further claims regarding age as a factor for variation. The terms *atchin tan* and *ker* represented two variants with two distinct Romani meanings. The term *atchin tan* was culturally symbolic and a reference to camping and stopping during travelling, Duncan and Jake's elicitation indicating a commonality of use of terms that were cultural meaningful or considered symbolic.

The use of *ker* and *kenner* translated as 'house' in European Romani and Cant respectively. The variant *ken* was found in the Derbyshire Travellers' recordings (Dawson, 1989-91: in Robert Dawson's Special Collection) whilst varieties of *ker* (*kier, keir, kair, kir, kar, Kerri)* were found amongst the archival materials (see Table 5.34 for archive details), with other close variants for 'house' (*gur, kaer, kar, ken, kenna, kir, korri, tam, tan*) being found listed in the Manchester corpus (2021).

With respect to information on background, outlined in the introductory section of chapter 4 and participant information in section 3.3.1, one interesting variant was the use of Graham's *ken,* showing borrowing and merging between Cant and Romani. Graham did not complete the ID measures (section 3.7) although his personal history was discussed. Background information was, therefore, less complete (see limitations 3.8.2) although Graham discussed his Roma identity as a part of his self-portrayal. Social network density, whose complexity might have been a factor influencing crossing and merging of what has been regarded as two distinguishable varieties used within the Traveller and Roma communities in England, and broadly found within the British Isles and Ireland (see limitation in terms of methodology section 3.8.2) . Clinton and Ben used *ker* , with both reporting close relationships with their fathers, and, further, identified openly as Romani. That might reflect greater contact, and, therefore, more proficiency, with potential for greater network ties to speakers of Romani. The term might be lexically less frequent and with less cultural status, therefore, greater exposure would be necessary for continuation and maintenance of the term. Network ties might also explain the use of Cant terms, such as, *kenner* as loose and multiplex social network ties (Milroy, 1980) or cultural interactions (White, 2008; Diehl, 2018) facilitates linguistic innovation. The use of *atching tan* for 'house' or 'home' represented a desire to associate a place of residence with a cultural association, along with a metaphorical notion of travelling.

The term 'here' was the only deictic variable within that SRN. The deictic variant was produced by Clinton, Jake and Ben. Jake and Ben used similar variants as Jake produced *aki* and Ben *akai*. Those variants were discernible by the addition of the vowel /ae/ with either a diphthong

for Jake or a monophthong, with a short final vowel for Jake's realisation. Those lexical items were orthographic realisations and phonological comparisons, based on speculations from written production. Clinton also produced *ochiakai /atʃiækæi/* during the interview which was also comparable. The final vowel was realised as a diphthong, with an initial prefixed *ochi /atʃi/.* There was a self-correction in excerpt 5.1.1 below.

Excerpt 5.1.1

1 Clinton ...*oh* ***otch aka****...*

2 ...*that means here* ***otch*** *means here* ***otch*** *means stop* ***okai*** *means here yeah…*

3 ...***ochi kai*** *that means you stop as well*

In the excerpt above, Clinton restated the meaning *otch* to mean stop as opposed to 'here'. The term *ochi* was later referenced in the excerpt as a term or command meaning 'stop'. That represented a frequent use of that variant, and was less likely to be parsed or used as a separate lexical unit, hence the repair of the definition. The confusion might also correspond with regional use of 'here' for 'stop' as in: 'here, what do you think about this?'.

Variants found within the archive data for 'here' are shown in Table 5.4. There was a degree of variation within the project data for that variable and Clinton and Ben's variants aligned with the Derbyshire lists collected by Dawson (2002). However, there existed considerable orthographic variation representing potential phonological variation within the archive and corpus data. Clinton's production of *Ochiakai /atʃiækæi/* was not transcribed in the archive or Manchester Corpus data. The transcription was assumed to represent a short, closed back vowel, which was potentially a unique phenomenon .

The variable 'town' elicited responses from Clinton, Jake, and Ben. *Gaff*, with the phonological variant *gav /gæv/* were produced by Clinton, and Jake and Ben respectively. *Gav*, *gab,* and the closed variant *garv* are found in the archive material shown in Table 5.4. Similar variants also appeared: *gov* (Sampson, 1891), *and gavas* (Dawson, 20thC: in Robert Dawson's Special Collection). *Gav* is defined as 'village' in European Romani (Manchester Corpus, 2021). The variant *leve* was also produced by Jake, and was not recorded as village or town in the archive or corpus. It might, consequently, represent a geographically and in-group variant.

Other variants recorded in the semantic category of 'place' were *drum* for Duncan, and Jake with the orthographic and phonological variant *drom /dɹam/* also being listed by Jake. The term

was European Romani and meant 'road' or 'way'(Manchester corpus, 2020). Whilst Duncan was less productive with lexis overall, interestingly, he produced the word for 'road' without hesitation, illustrated in the excerpt below. During the discussion about lexis Duncan embellished his with references to several notions that connected across each of the informants' interviews shown in excerpt 5.1.2

Excerpt 5.1.2

1 Duncan: *...alot of Gypsy people there that know alot about horses there as well you know*

2 *it's an old an old you know custom an old stuff that they know is unbelievable*

3 *about horses I mean I mean…*

4 *...I mean there's a family there called the Bullman's from north of England they*

5 *trade horse for Prince Philip…*

6 *...there that's one for you and they're Gypsy people as well…okay* ***drum***

7 *is road ...****drum drum***

As previously mentioned in excerpt 4.1.5 in chapter 4, reference to horses was a thematic notion referenced across the participants' interviews. That was expanded with the association of aristocracy or nobility in the discussion in 5.1.2 where Duncan associated a notion of horse related activity, and, in particular, with reference historical events involving associations between the Anglo-Romani and aristocracy. That placed the importance and status of horse expertise as he associated that as a cultural characteristic of Anglo-Romani by referencing: *'they trade horse for Prince Philip*' outlining his affective stance towards the practices of horse related knowledge.

Whilst Duncan produced fewer lexical variants, his discussion of culturally relevant themes was significant. His discussion of variation with significant aspects of cultural transmission were relative to a self-perceived identity. In contrast, although more productive, Clinton produced no variant for 'road', after prompting, shown in excerpt 5.1.3, however, Duncan did. Therefore, whilst some terms might seem to have cultural significance, that was not always a factor in terms of the productivity or maintenance of specific lexis between the individuals within this study.

Excerpt 5.1.3

1 Interviewer: *places road just a road for travelling along*

2 Clinton: *no just a road again*

Significantly, Ben also did not produce a variant for 'road' which might be related to personal circumstances and contexts of acquisition or attrition. Factors that influenced variation and productivity were further explored with reference to biographical, contextual and linguistics factors. Jake produced *longo drum* and *longo,* a variant of English 'long' (Manchester Corpus , 2021). Jake, also, produced the only word for a place name *Rushni Tem* meaning Lincolnshire although other place names have been recorded in the archival data. Notably, *Rushni Tem* did not correspond with other archival material and represented a unique recording. However, the Romani word *tulo-mas-them* was recorded by Matras (2010a) meaning 'bacon country' in reference to Lincolnshire's farming association. *Tem* was European Romani meaning country, and was found used as a suffix for some place names in Anglo-Romani. An example was *wudrusgavtem* meaning bed, (*vodros*) village, (*gav*) country (*tem*) (Manchester Corpus, 2021).

The term for 'Outside' was listed in the SRN and Jake produced the variant *aurial /aɹiəl/*. This was likely an orthographic and phonological variant of *avral /avɹiəl/* meaning 'Outside' in European Romani (Romlex lexical database, 2020). This represented a lower frequency item, illustrating the tendency for Jake, and more generally Clinton and Ben to produce a significantly greater number of non-standard variants. Background factors for Jake included being from an earlier generation, with dense network ties (see Table 4.5, and 3.8.2 for limitations), and an active participation in cultural activity.

The variables for 'caravan', 'toilet', and 'Lincolnshire' were not listed in the SRNs and represented spontaneous elicitations.

5.1.3 Money and Numbers

There were a total of 24 variants for that semantic category: 'Money and Numbers'. This was a comparatively productive category although less so than 'places' or 'vocation'. Notably, many of the variants differed in terms of orthography and were transcribed by the informant in the case of Jake and Ben. Those orthographic variants represented phonological variations in certain cases (see section 3.5 for clarification of this distinction), and this was discussed where relevant. Below, Table 5.5 illustrated the lexical variants elicited for the semantic category of Money and Numbers.

Table 5.5 SRN: The Outside World 2: Money and Numbers

Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Number one				*Bar*	*Yek* DV1,D1, V1, F1, S1, B2, SC1, L1, S2, LSP	
Number two				*Doi* W2, DV2, N1, R2,G, EA, S3	*Doui* W2, DV2, N1, R2,G, EA, S3	
Gold					*Shuvler*	*Sunnikai/ galaipe* F1, S1, B2, SC1, MC, R, MG, S3, N1/
Half					*Pas*	
A pound				*Luvver/ Bar* /DV1, DV2, LSP, B2	*Baar/bar* /DV1, DV2, LSP, B2	
Cash				*Luvver*	*Lovvel*	*Lov*

Money	*Greid/ Gorid* DV4/DV4	*Luvver/Wonger* MC, DV2, SC1, EA, V1, F1, B2, S1,SC1, L1 /SC1, MC	*Panch/ Luvva/ Grad* /MC,DV2, SC1,EA,V1, F1,B2,S1,SC1, L1/		*Luva* MC, DV2, SC1, EA,V1, F1,B2, S1,SC1, L1	
Fiver (Five pounds)				*Flag*		
A fortune					*Dosta Luver*	

Table 5.6 Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

The variable with the most variants was 'money'. Both 'money' and 'cash' were included in the SRN, with one informant responding with distinct variants between those two semantic

concepts. Jake responded with *lovvel* for 'cash', and *luva* for 'money'. That distinction was not found in archival or corpus material. Within the archival material, 'cash' and 'money' were recorded synonymously, and within the Manchester Corpus (2021) the distinction might be put forward as there were no results for 'cash' (see Table 5.5 for this distinction found within the data). There were numerous orthographic variants representing potential phonological distinctions within the archive and corpus data. However, the variant for 'cash' *lovvel* was not found within the archive as a term used for 'money' and represented a variant unrecorded. The term *luva* used by Jake for money was found in the archival material (Croft and Smart, 1875) along with similar variants, and these have been provided here to example the comparative forms found within the archive sets. These were: *lova* (Dawson, 20thC); *luvva* (Dawson, 2002; East Anglian, 1929) *luvna* (various sources, 20thC); *lovo* (Fox, 19thC); *luvvo,* (Borrow, 1874; Sanderson, 19thC) *loovoo*, *luvo*, *wongar*, , (Sanderson, 19thC); *luva*, *vongar*, *vangar* (Sanderson, 19thC); *angar*, *angar*, *wonger* (Croft and Smart, 1875); and *love*, *vanga* (Lucas, 20thC). Within the Manchester Corpus also considerable variation had been recorded; *larvo (*Norwood, 19thC); *lauwi*, *lovva*, *lowvi*, *lulla*, *luv*, *luvvay*, *luvverd*, *luvvi*, *vonga*, *wonga* (Manchester Corpus , 20thC); *lovvo* (Irvine, 19thC); *luvvo* (MacGowen,1996; Russel, 1916); and *luvvu* (Roberts, 1912).

The variant *luvva* was used by Clinton, which corresponded exactly with the Derbyshire word lists (Dawson, 2002). This suggested a regionally distinct variant. Graham, Duncan, Clinton and Jake produced comparable variants in terms of orthography, representing an initial mid-closed back vowel and final open back vowel in /luvə/. Phonological comparisons here have been based on orthographic realisation set out by the dialect society referenced by Dawson, and used in his transcriptions (2002). For Clinton and Ben, the term for 'money' and 'cash' were recorded as synonymous, and documented as 'cash' which appeared to be more consistently the case across archival sources. Duncan used the three variants, *panch, luvva* and *grade*. The variant *panch* might represent an orthographic variant of *poshas* meaning 'half' which is recorded as being used for 'money' (Manchester Corpus, 2021).

A point of interest was the use of *grade* by both Judy and Duncan. Judy self-identified as Irish Traveller, whilst Duncan as a Romani with Irish Traveller heritage. Below in excerpt 5.1.4 is Duncan described details about this cultural background.

Excerpt 5.1.4

1 Duncan *...yeah well we say money I told you the other day money was* ***luvver*** *yeah okay*

*so if it was big money we would say **boot luvver**…*

*…yeah yeah because my mother's Irish and where we would say instead of saying my father would say **luvver** which is for money my mother would say **grade**…*

*…**grade** that's an Irish word it's Irish Cant and English **Romanesc**…*

…I have both yeah and I know alot from the Irish community as well because I have alot of family amongst the Irish community so then I had both yeah alot of English Travellers don't have that I've got both and there's alot of…

Duncan reported that his mother used the term *grade*. This explanation was significant in that it distinguished between Romani and Irish Pavvee Cant, which were often characterized as being merged. Duncan described this distinction as overt as well as being culturally distinct in terms of his family's heritage, whilst *grade* was expressed to be both Irish Cant and *Romanesc*. In excerpt 5.1.4, Duncan made reference to '*the Irish community'* and his heritage, was referred to with the referent '*English Travellers'.* The use of both lexical variants were associated with, and important, in his own background in'*…grade that's an Irish word it's Irish Cant and English Romanesc…*' and established his affective stance towards his use and understanding of these terms in'*…I have both yeah and I know alot from the Irish community as well'* concluding '*I've got both and there's alot of…'.* Within that narrative, Duncan discussed his own heritage as '*English Travellers'* having an additional dynamic to his background in '*I've got both'*.

Judy self-identified as an Irish Traveller (Table 3.0), who resided in Leicestershire. She uses the term *gried and gorid* for money. Both terms are recorded by Dawson (2011) from word lists based on a collection from a number of historical sources collected from self-ascribed speakers of Pavee Cant, showing how the term for 'money' was found across sources within the archive and further discussed within the data presented here.

The data revealed certain lexis that were perceived by self-ascribed speakers as distinct between Irish Traveller Cant and Anglo-Romani, and were persistent in terms of their use as markers for ethnic categorisation. In addition, and as previously highlighted, examples from the interview data can be considered in terms of a folk-linguistic accounts of lexis and etymology. In Excerpt 5.1.5, a folk-linguistic account was based around the sources of the word

for ‘fight’. Here, the term was described as being used for ‘pay’ and deriving from the Spanish for ‘fight’ and based around a discussion of the word for ‘money’.

Excerpt 5.1.5

1 Duncan: *…he [Matras] realised because he was a man that could speak that many*
2 *different languages he realises you know but see I think me personally the word*
3 ***pagge**r is a Spanish word for pay so maybe maybe is it because we came to this*
4 *country and maybe travelled through Spain and then picked the way to fight for*
5 *your money because we use fight for pay for a fight **pagger** is a word for fight*
6 *in Spain **pagger** means pay I can speak Spanish…*
7 *… maybe you had to fight for your money you know when you have done your work…*

In this excerpt, Duncan made the association of heritage and ancestry of Anglo-Romani links to European migration by suggesting the word *pagger* relates to the Spanish word with ‘*because we use fight for pay for a fight pagger is a word for fight in Spain pagger’.* Whilst *phag* links European Romani as an origin of the word *pagger*, the notion of a Spanish origin describes cultural and personally relatable attributes for an origin for this word as relative to someone who had to work hard or ‘fight’ to make a living. The term for fight is recorded as *pagger/pogger* in the Manchester Corpus (2021) although there is no record of its use in this context. In the Manchester Corpus the term is recorded as a cognate with European Romani *phag*(er) meaning ‘break’. However, the term for ‘pay’ in Spanish is ‘pager’ representing a separate and plausible origin.

The semantic category for ‘Money and Numbers’ elicited few variants for numbers. This suggested that whilst some of the participants were exposed to Romani there was less salience for this category of lexical terms. Possibly, terms such as for money and numbers would have been less important as part of an in-group lexicon. The numbers ‘one’ and ‘two’ are the only numerical terms elicited as seen in Table 5.5. The terms for ‘one’ and ‘two’ appeared across two participants Clinton and Jake as *doi* and *doui* respectively. Phonological distinctions for the variants produced by Jake are speculative as they are based on judgement for frequent orthographic realisations (see section 3.8.2). Numbers are a feature in many of the recorded archives and variations of *doi* and *doui* are common; *dui, doui, due, dui, dooe,* (Winstedt, 1948;

Dawson 2002; Norwood 19thC; Russell, 1916; Roberts', 1912; Grosvenor, 1908; No Author, 1929; Sampson, 1911). *Yek* for 'one' appeared across several sources from the archive material (Dawson, 1959-69; Dawson, 20th; Various sources, 20thC; Fox, 19th C; Sanderson, 19thC; Borrow, 1874; Croft and Smart, 1875; Lucas, 20thC; Sampson, 1891), *yak* (Leeds Special Collection, 1979). No results appeared within the Manchester Corpus for 'one' whilst the use of *bar* for 'one' by Clinton is elicited. Whilst numbers are a feature that can be traced back to Indo-European roots and *bar* not an exception, it is assumed that the word *bar* is less direct in terms of etymology. In excerpt 5.1.6 below, Clinton explains the use of lexis used for numbers five, one two and half a shilling and his use of use of *bar* for 'one'

Excerpt 5.1.6

1 Clinton: *...yeah whatever like **flag** means fiver **bar** means one **doer** means two pounds*

2 *half a **bar** means ten shilling that used to be pennies*

3 *and shillings didn't they I did do anyway…*

In this example, Clinton explained the use of the term *bar* contextualised within discussion of currency denomination. The context highlights the use of the term *bar* extended to 'one' although it is not clear that this example use of *bar* used for numerical 'one' outside of this context are suggested from this example.

The term *bar* and similar variants were found across archive material and the online corpora. The term is used for currency meaning one pound sterling (Dawson, 1959-69; Dawson, 2002; Leeds Special Collection, 1969; Borrow, 1874). The cognates *balans* (various sources, 20thC) and *ballans* (Sanderson, 19thC) are recorded as used for both 'money' and 'weight'. Whilst *balanser* (Way, 19thC) *bas, base* (Borrow, 1874) and *balanser*, *balans* (Croft and Smart, 1875) are also recorded as used for 'a pound'. The word for 'stone' in European Romani is *bar* (Manchester Corpus, 2021) and a clear suggested origin for *bar* used for a pound sterling. The term for 'stone' would have used as a value term relating to weight of an item. Historical use of stone with monetary value include the Rai stone of Micronesia (Leonard 1993) for example.

Clinton also spontaneously produced the term *flag* for the currency denomination for five pounds. Whilst *bar* appeared across several sources and its origin suggested in the Manchester Corpus as European Romani *bar* meaning 'stone' or 'rock', *flag* does not appear within the archive material. Interestingly, the word for 'five' in European Romani is /pandz/ and recorded as *panch* and *ponch bar* in Anglo-Romani (Manchester Corpus, 2021). Duncan reports *panch*

for money which may represent an extension of the European Romani word for ‘five’ in the context of money for Duncan.

The variant *shuvler* recorded by Jake and *galaipe* for Ben did not appear in the online corpus or archive material. These represent unrecorded alongside a number of others as highlighted through these chapters. The term *sunnikai* recorded for Ben for ‘gold’ appears across several archive sources; *sunakie* (Fox, 19th C) *sunika*, *sunika* (Sanderson, 19thC) *sonnaki* (Borrow, 1874) *soonakei* (Croft and Smart, 1875) *sonakai, sonnakai, sunnakai, sunnather, sunnikai* (Manchester Corpus, 2021), *sonnokai* (Russell, 1916) *sonnokoi* (MacGowan, 1996) *suhaki* (Sampson, 1911) and *sunakai* (Norwood, 19thC). The cognate term *sonakaj* is recorded as European Romani for ‘gold’ (Manchester Corpus , 2021).

Variants were limited to one variant in the case of ‘half’ and ‘fiver’ and ‘fortune’ and were spontaneous productions not featured on the SRNs.

5.1.4 Time

The semantic category ‘time’ shown in Table 5.7 shows Clinton produced three deictic variants related to time during spontaneous conversation. The five other participants were not prompted with these variables, however. Table 4.5 displays the variables for ‘time’ with four variants elicited by Clinton.

Table 5.7 SRN: The Outside World 2: Time

Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Night				*Rati* MC, N1, V1, S1, B2, L1, IV, R2, F1, DV2, G, FS, W1, SC1		
Tonight				*Torate* MC,DV2,G		
Morning				*Sorla* S1,DV2,G,W1, SC1,L1,DV1,MC,EA, R2		

Day, **Today				*Divas* MC,S3,G,F1,S1,W1,B2,SC1,DV3,L1,S2		

Table 5.8 Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

Clinton produced the variant /*ra:ti* /for variable night and /*sarla/* for morning shown in the interview excerpt below. The words correspond with European Romani *rati* and *tesarla* for 'night' and 'morning respectively' (Manchester Corpus, 2021). For comparison, similar variants for 'night' were found in the archive material and coded within Table 5.7 which are; *rati* (Norwood, 19thC; various sources, 20thC; Sanderson, 19thC; Borrow, 1874; Lucas, 20thC; Irvine, 19thC; Roberts', 1912), *rattie* (Fox, 19th C; Dawson, 2002; Grosvernor, 1908; Fox and Samuel, 1926; Sampson, 1911), *rarde* (Way, 20thC; Borrow, 1874) *raati* (Croft and Smart, 1875), *rate* (Lucas, 20thC) *darky* (Leeds Special Collection).

Clinton also produced the variant /tə'ra:ti:/. This variant was found in the Manchester Corpus and archive data; *torati* (Dawson, 2002; Manchester Corpus, 2021), *araati* (Dawson, 2002),

kerat (Grosvernor, 1908). The prefix /tə/ is considered as English preposition 'to' (Manchester Corpus, 2021).

Variants for /*sarla/* also appeared in the archive material*; Sarla, Salo, Sawlo* (Sanderson, 19thC; Dawson, 2002; Grosvernor, 1908), *Sala, Sorlo* (Way, 20thC), *Saula, Saala* (Croft and Smart, 1875), *Sarla* (Lucas, 20thC) *Sarle* (Dawson, 1959-69) and the Manchester Corpus; *sala, sora, sowla, sawla* (Grosvernor, 1908; No Author, 1929), *sowla* (Roberts', 1912).

The items *rati* and *sawla* were elicited towards the final stage of the interview. In excerpt 5.1.6 Clinton asked the rhetorical question '*what does that mean?*' to create anticipation. Clinton shows confidence in his propositional stance, demonstrating knowledge and authenticating his information and his understanding of dialect within the interaction. This positioning was evident across interactions, outlining the principles proposed by Bucholtz and Hall (2005).

Excerpt 5.1.6

1 Clinton *...well let me ask you tell you this you've got* ***rarty*** *what does that mean…*

2 *...it means night*

4 *...**rarty** means night… or you've got it*

5 *...**sawla** mean morning…good **sawla***

In the following excerpt in 5.1.7 Clinton corrected the interviewer in the pronunciation of the phrase a *peeve torate* with '*no a* ***peeve torate'*** during an explanation of the meaning of the phrase and lexical variant *torate*. This epistemological stance highlighted the role of Clinton as authenticator. The emphasis below is given to phonological/phonetic realisation.

Excerpt 5.1.7

1 Clinton: am having a *peeve torate* that means I having a having a beer tonight…

2 Interviewer: a ***peeve torate***

3 Clinton: *no a **peeve torate** that means I'm having a drink tonight…*

The variant /divəs/ for 'day' was produced spontaneously in conversation by Clinton. The term *dives* is European Romani meaning 'day' and was found in the Manchester Corpus (2021 as well as the archive sources; *devvus* (Sampson, 1911) *divvus* (Manchester Corpus, 2021), *kedivvus* (Grosvernor, 1908), *tedivvus* (Manchester Corpus, 2021), *todivvus* (Russell, 1916), *dives* (Bright, 19thC), *divesəs* (Various Sources, 20thC), *devus* (Fox, 19th C), *divvus*

(Sanderson, 19thC), *divvus* (Way, 20thC; Borrow, 1874; Croft and Smart, 1875; Dawson Special Collection) *dives* (Croft and Smart, 1875), *divas*, *devel* (Lucas, 20thC), *kedivez*, *kova*, *divvus*, *bedivvus*, *kedives*, *kee devus* (Dawson, Special Collection), *dives* (Sampson, 1891).

Highlighting an inflectional feature of Anglo-Romani, in excerpt 5.1.8 below, Clinton contextualised the use of *divas.*

Excerpt 5.1.8

1 Clinton: ***kushty divas*** *today that means its nice today*

This illustrated a multi-word phrase whilst in terms of inflectional characteristic for Anglo-Romani there is a contrast of integration of English 'to' in informant's production of *torati* shown in Table 5.1, whilst *divus* does not take the prefix 'to' within the usage for Clinton.

5.2 Conclusion

In conclusion to the above analysis and discussion, several salient features were evident in terms of lexical variation and identity within the interactional data. In terms of lexical variation outlined here, there was a pattern of variation that was aligned with the archive data sets used for the purpose of comparison. However, there was notably a degree of variation that cannot be explained without reference to individual and external factors of influence. In sum of the lexical variation, there were a number of traits. Irish Pavee Cant and Anglo-Romani continue to show as distinguished sources of lexical variation and were utilised significantly as separate entities as attested by the participants within the study and described with example within the chapter. Individual factors accounted for the degree of inflection, innovation and productivity, such as the lexical repertoires of the participants such as Jake and Ben's usage of certain lexical variants. Identity Index Scores were discussed as showing attitudinal variation as an influencing factor, as was exposure at a younger age. A number of terms were speculated to have regional or diatopical (locality) relation. Whilst age did not present as a significant factor, the principle that locality has an influence on the lexical form used is proposed based on a number of examples. In addition, a number of original attestations were also discussed with examples concluding two points: that further research is necessary to ascertain the extent of lexical variation within the communities of practice, and that the methodology used for this project is suitable for attaining unattested lexical data from a minority language speaking community. In addition, there was a suggestion of gendering as an influence on lexical retention, and use and meaning of gender neutral terms needs further attention and again,

further work would be necessary to determine this as a factor for variation within Anglo-Romani and Iris Pavee Cant.

In terms of interaction, meta-linguistic discussion was seen utilised as an indexical act of identity practice and these interactions are seen to enable the association of cultural identity within discourse. Several occasions and throughout a number of discussions, the origin and etymology was used as a means for cultural reference and ethnically positioning the speaker. Within several interactional episodes the use of pronouns and reference to lexis as part of the community established a cultural alignment with the variant terms. In addition, reference to a number of cultural associations also established a social position as culturally indexed. These included a reference to authority and attitude and stance towards this authority. This was achieved with reference to hierarchy, status and acts within hypothetical scenario. These interactional analyses were framed within observations of interaction that show identity as an inter-subjective entity, emergent through a process of identity construction. This was brought out through observation of indexical reference through topic choice, saliency and culturally aligned phenomena, as well as the meta-linguistic discussions themselves around lexical terms. The following chapter picks up on these themes within the lexical data focussed around the semantic category ‘Feelings, Actions and States’.

Chapter 6 Analysis: Feelings, Actions and States

6.1 Semantic category three: Feelings, Actions and States

The SRN 'Feelings, Actions and States' were divided into four categories. These categories are 'Doing things', 'Feelings', 'States', and 'People' (see section 3.5 for a detailed description and outline of the development and selection process of this SRN). There were a total of 213 variants recorded for the SRN 'Feelings, Actions and States' from a core of 55 variables. Lexical variation elicited from the SRN 'Feelings, Actions and States' were analysed with reference to Identity Score Index, Identity Questionnaire, and Visual Analogue Scales and the Interview data. Identities through social practice were further presented from the principles of the framework outlined in section 3.8 relating to Bucholtz and Hall's paper (2005) further developing the analysis of discussion work from chapters 4 and 5. Chapter 6 further develops the representation of each participant's dialect repertoire, situated within background, context of use and historical context, regional patterns and unrecorded attestations. Analysis of the archive and corpus data allowed comparable and contrastive accounts of non-standard variant use. In addition, an analysis of social meaning in practice outlined the negotiation of identity through social means, supporting an understanding of identity practices understood through linguistics and more specifically, lexical choice.

6.1.1 Doing things

The most productive subcategorisation was 'Doing things' and analysed first. As previously discussed, variables with a higher degree of productivity are analysed first and those with least productivity last. As with all SRNs, the codes below each variant correlate to archive sources. The key to these codes is given in Table 6.1. There were two variables that elicited variants from all the participants, and these were 'go', and 'look' shown is Table 6.0.

Table 6.0: SRN: Feelings, Actions and States: **Doing things**

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Piss, Urinate		*Gelling to the* *Muttering* *Kenner* DV2, D1, S1, W1, SC1, L1				

Go, Going	*Crush, Shift* LSP,DV4	*Gel,Gelling *Gel* DV1,DV2	*Gel, Crush 'crush her sublik'* DV1,DV2/LSP	*Gel* DV1,DV2	*Jel* DV1,DV2	*Jal, jol* D1,N1 ,S1,B2 ,SC1, DV2, EA, Ev/SC 1
Come			*Crush* DV4			
Work		*Booty* DV2, S2, S1, V2, W1, B2, SC1		*Buttei* DV2, S2, S1, V2, W1, B2, SC1		
Give		*Dell* DV2, R, EA, SC1, DV1, LSP, D1,T1, S1, W1				
Run			*Gel/Missly* MC, B2,L1		*Nash, Jell DV2,W2, R2/MC, B2,L1*	*Prast a* MC, R, IV, DV3
Do/*Make				*Buttei* See 'Work'	*Kel (to do)* DV2, S1, SC1	**Ker/k erava* MC, B2, SC1
Take			*Puv*			
Stay			*atch (here)* R, EA, S1, B2, SC1, L1			

Look	*Sorni (Sornin de fin – Sornin de byoua, Sorni de Pavee)* DV4	*Dik* R, DV2, EA, DV1, D1, V1, F1, S1, W1, B2, SC1, L1, DV3	*Yag* DV2	*Dik* R, DV2, EA, DV1, D1, V1, F1, S1, W1, B2, SC1, L1, DV3	*Dik* R, DV2, EA, DV1, D1, V1, F1, S1, W1, B2, SC1, L1, DV3	*Dik, dita* R, DV2, EA, DV1, D1, V1, F1, S1, W1, B2, SC1, L1, DV3/
Fight/*beat	**Korob (beat)* MC, LSP, DV4, DV5		*Paga* MC	*Kor* F1, S1,W1,B2 ,SC1, L1, S2	*Kor* F1, S1,W1,B 2,SC1, L1, S2	*Cor* F1, S1,W1 ,B2,SC 1, L1, S2
Raid		*Steaming*				
Sell/Hawkin g		*Monging* MC, SC1				
Camp/Stop over	*Pull on*	*Atchin Tan* MC, SC1, S1, W1, B2, L1		*Otchikai torata* S1, SC1, W1, B2. L1	*Atchin Tan* MC, SC1, S1, W1, B2, L1	*Artchi ntan, tan* MC, SC1, S1, W1, B2, L1
Get (for free)/Steal				*latchet (for nickcess)*	*Kin*	*Lela/ astav* MC, LSP, S1,V2,

				R2, MC, (MC, DV2)		W1, B2, SC1, L1, S2, DV2/
To sleep,			*Suttty* DV2, R2, EA, MC, DV1, N1, V1, S1, B2,			
Curse				*Amock*	*Chinger* DV2,	*Waffad o Lav, Shava ni Pukar as* MC, DV2
Tell				*Boore*	*Pucker* MC, DV2, V1, B2, S1, W1, B2, SC1	*Pen, Puka* S2, N1, V1, DV1, R/ MC, DV2, V1, B2, S1, W1, B2, SC1
Talk		*Rokker* MC, IV, R, DV2, G, DV1,	*Rokker* MC, IV, R, DV2, G, DV1,	*Rokkerin* MC, IV, R, DV2,	*Rokker* MC, IV, R, DV2,	*Rokka/ Puka*

		D1, S1, W1, SC1, L1	D1, S1, W1, SC1, L1	G, DV1, D1, S1, W1, SC1, L1	G, DV1, D1, S1, W1, SC1, L1	MC, IV, R, DV2, G, DV1, D1, S1, W1, SC1, L1/MC, DV2, V1, B2
Read fortune		*Dukkering,Ducker* MC				
Know				*gen* MC,DV2, DV3, V1, S1, W1, B2, L1,		*jin* DV2, R,R2, D1,S1, W1, L1
Burn/Fire			*Yog* DV2, MC		*Yog/otcher (to burn)* DV2, MC/DV2, R2,	
Stop	*Stal, Shtaul, Stul,* MC, DV4, DV1	*Atchin* MC, S1, W1, B2, SC1, L1, DV3,		*Ochika* MC, S1, W1, B2, SC1, L1, DV3	*Arch, Brishindo* MC, S1, W1, B2, SC1, L1, DV3/DV2	*Artch* MC, S1, W1, B2, SC1, L1, DV3

Table 6.1: Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

Table 6.0 showed Graham, Duncan, Clinton, and Jake producd the variant *gel/jel* for variable 'go'. In relation to the archive sources the variant was found limited to vocabulary lists in the East Midlands (Dawson, 2002, 2011). This suggested the variation is regional (see section 3.8.2 for discussion of methodological limitations). Ben's use of the variant *jal* or *jol* was distinct from *gel*. Ben described his background as being from New Zealand with a Welsh Romani heritage (see section 3.3.1 and Tables 4.1) influenced by his with Welsh heritage through his Romani speaking father. He also describes himself as an active speaker within an international network of a wider Romani speaking network. Judy used *crush* and *shift* and this is comparable to Duncan who described himself as having an Irish Traveller heritage (Table 3.0 and Figure 4.3) who also attests *crush*. *Crush* is found in the Irish Traveller Pavee Cant word list of Dawson (2011) and aligns with both Judy and Duncan's self-attested Irish Traveller background. The term *shift* was used by Judy and is vernacular English 'to move' and is not found within Irish Traveller dialect archive sources. For Ben, both *jal* and *jol* aligned with variants found within the archive sources. Most significantly, these vary from those elicited from the participant within the East Midlands. The cognate term *džal(a)* means 'go' in European Romani (Manchester Project, 2021) and is a cognate for *gel, jel*, *jal* and *jol.* Excerpt 6.0 shows Clinton using the term *gel* in context and with cultural reference to horse.

Excerpt 6.0

1 Clinton: *if you were driving your horse on you'd say* ***gel***

In excerpt 6.0 Clinton contextualised the use of *gel*. The reference to horse related activities as a cultural index is discussed in chapter 5 (see discussion relating to excerpt 4.1.5). Activities that are symbolic or indexical to community identity is a pattern found within coal mining communities also. Those that identify as 'Miners' were found to utilise more regional variants when talking about their trade (Devlin, 2014). A comparable pattern was suggested with participants within this study who self-ascribe to Romani and Irish Traveller. This association is made on several occasions and is discussed as a significant cultural dimension in terms of Clinton and specifically in relation to use of Anglo-Romani dialect variants as discussed in excerpt 4.1.5. Whilst terms such as 'gee' used for animal life are associated with regional variation and specifically the East Midlands (Braber, 2018: 128), the verb *gel* is used by Clinton arguably as symbolic to his description of his own identity as being part of a wider Anglo-Romani community.

The variable 'look' also elicited variants across all participants. The term *dik* was used by Graham, Clinton, Jake, and Ben. The variant *dik* was found across several archive sources of Anglo-Romani word lists and a there was uniformity to the realisation of this variant. The European Romani term *dikh* (Manchester Project, 2021) meaning 'see' is a comparable cognate for this dialect variant. Duncan produced *yag* however, and this is also found in the archive lists, although regional to the East Midlands (Dawson, 2002). The variant is idiomatic in terms of its metaphorical association in relation to its etymological cognate *jakh* which means 'eye' in European Romani (Manchester Project, 2021). Lexical innovations are seen as a pattern historically within Romani and within the present data (sections 2.2.1.2, and discussion of excerpt 4.8 and 4,9). However, lexical innovation of Romani words including semantic calques such as *nevo foro* 'new city' for Neustandt in German Sinti and semantic enrichment is under explored (Scala, 2020). Variants such as this show examples of Anglo-Romani being adapted and how lexical items are used. Ben also produced the variant *dita /ditə/*. Although not attested in any of the Anglo-Romani word lists within the archive, there is a clear cognate between the European Romani form and represents a phonological variation rather than lexical or etymological variation.

For Judy, *sorni* was elicited which is found within Irish Pavee Cant dialect sources as recorded by Dawson (2011) attested by self-ascribed speakers of Pavee Cant (see section 3.8.2 for

discussion of limitations). Excerpt 3.7 showed from field notes the production of *sorni* meaning 'look' and attested by Judy. The phrases from excerpt 6.1 translate as 'look at the boy', 'look at the girl' and 'look at the Irish Traveller' respectively. These represent some of the few examples of contextualised variants notated from Judy and show lexis embedded within phrasal syntax.

Excerpt 6.1

1 Judy: ***Sorni** de **fin – Sorni** de **byoua, Sorni** de **Pavee***

Excerpt 6.2, shows Clinton using the term '*dik*', explaining its use in context embedded within a representative Anglo-Romani sentence.

Excerpt 6.2

1 Clinton: *no you'd say **dick** to the **gearers toggies** he's **kelt** up **kushty toggies***

2 *means clothes*

3 Sister-in-law: ***dick** means look…*

4 Clinton: *it ain't Townsville slang it's come from Gypsy people*

5 *but some of the words is Gypsy people's*

6 *and some of it what they've added on you know what I mean*

In this exchange (Excerpt 6.2) both Clinton and his sister-in-law explained the meaning of terms for the interlocutors. The use of Romani dialect is discussed by Clinton and sister-in-law as used to convey meaning between family members and friends as a covert or 'secret language' as market traders. 'Secret languages' referring to Halliday's (1975) idea of an in-group code can be associated with communication between individuals who are part of an in-group such as traders as they may need to talk covertly and discreetly liaise. Whilst Anglo-Romani and Pavee Cant may be used to talk covertly it is not a defining characteristics of the language or dialects (Matras, 2010: 106). In excerpt 6.2 the conversation turned to the use of some Romani words being defined as 'slang' in the local area. This status was challenged and dismissed by Clinton in the exchange. Changes to the use of lexis are acknowledged as a result of community innovations (Matras, 2010) and Clinton acknowledges this when he stated '*and some of it what they've added on you know what I mean'* suggesting a distinction between his definition of 'slang' and Anglo-Romani dialect status. Whilst it may be used within a wider local community,

here Clinton was positioning its use within the 'Gypsy' community and claiming this lexis as belonging to the 'Gypsy' community and establishes this with use of the pronoun '*we*' and '*they've'* referring to the wider community. This affective position towards this ownership demonstrates Clinton's attitudinal stance towards Anglo-Romani as becoming part of a local variety although not being 'Townsville slang' but 'comes from Gypsy people'.

'Fight', 'camp', 'stop' and 'talk' elicit variants from five out six of the participants. Duncan was less productive in terms of elicited items for 'Doing things'. Figure 4.2 highlights comparable productivity between participants across all sematic categories and showed this as an overall trend for Duncan. This pattern was seen within this SRN. The variable 'fight' elicited *korob* from Judy. The variants found in the Manchester Anglo-Romani corpus and dictionary are not a direct match although *korben* is similar in terms of its comparative transcriptions. The European Romani cognate given is *kur* ('beat (up)') which is similar in form. Clinton, Jake, and Ben used the variant *kor* or *cor* which matched variants listed in Anglo-Romani archive sources. However, variant *korob* used by the participant self-ascribing as Irish Traveller is distinct from the Romani *kor*. The term *korob* did not match the variant used in the Leeds recording recorded by a participant self-ascribing as an Irish Traveller dialect speaker and more closely matches *korib* found in the Irish Traveller Pavee Cant word list of Dawson (2011). Either there are two distinct semantic roots for these variants, or meaning related to social practices have influenced this distinction in forms for a variant rooted in the same cognate. However, this is not attested elsewhere in the literature. Duncan used the variant *paga* for variable 'fight'. The variant is found in the Anglo-Romani corpus (2021) and means 'break' in European Romani (Manchester Project, 2021). Excerpt 6.3 below illustrates a folk-linguistic account as described by Albury (2011) and further by Rieder (2018) discussed in relation to excerpt 5.0, and centres around the premise of 'working hard' or 'fighting' to make a living.

Excerpt 6.3

1 Duncan: *he realised because he was a man that could speak that many different*

2 *lanaguages. He realises you know but see I think me personally the word*

3 ***pagger*** *is a Spanish word for 'pay' so maybe maybe is it because we came to*

4 *this country and maybe travelled through Spain and then picked the way to fight*

5 *for your money because we use fight for pay. For a fight pagger is a word for*

This account centres around the dialect being influenced by contact through migratory movement. Duncan used modality to build speculation of the etymology of the word '*pagger*' as he stated '*He (linguists/Matras) realises you know but see I think me personally the word **pagger** is a Spanish word for 'pay'…'.* In this example, Duncan's proposition explores the idea of an etymological source of the word *pagger* supported by his experience of Anglo-Romani and Spanish. This type of speculation is themed across his discussions and acts as contextualisation of an identity raised within the interviews.

'Camp' is realised as *atchin tan* by Graham, Jake and Ben. Either the variable is realised as *tan* or *atchintan* within the archive lists and translates as *ačh*- 'stay' and *than* is 'place' in European Romani (Manchester Project, 2021). Ben used both *atchintan* and *tan* and this is consistent with variant use found in the archive lists and is uniform across varieties of Anglo-Romani. Clinton, however, uses the term *otchikai torata* which translates into 'stop here' (*akaj* 'here' in European Romani). Clinton also included the word *torata* although this was not found amongst the vocabulary word lists within the archive used for 'camp/stop over'. Judy used an English variant *pull on* which may be characteristic of a mixing of English vernacular within her dialect repertoire.

The variable 'stop' elicited *atchin, arch* and *artch* from Graham, Jake and Ben respectively. The term is a cognate with European Romani *ačh* - 'stay'. The word for 'stop' and 'camp over' is an example of an Anglo-Romani dialect word that is used in more than one context and is a pattern of Romani language and patterns of extension that relate to language contact (Matras, 2010: 106). This may be an extension of the European Romani word ačh meaning 'stay'. Clinton also used *ochikai* for 'camp' as well as 'stop' and is a compound of *ochi* and *akai* (*akaj 'here' in* European Romani). Use of *ochikai* for 'camp' and 'stop', although referenced with archive variants, is unique to Clinton amongst the archive lists. Whilst archive codes are referenced, they are proximal (defined in section 3.5) rather than a direct match with Clinton's use. Jake also used the term *brishindo* which may be a variation on *besh* found in Dawson's Derbyshire lists (2002). This is a cognate of *beš* - 'sit' (Manchester, 2021). The forms found for this variant are not close matches orthographically or in terms of phonology within the archive. This then may represent an archaic form related to Anglo-Romani specific to Lincolnshire and the dialect branch of the Northern Romani dialects as these are isolated to the region. Chapter 2 outlines in detail the dialect branches (see section 2.2.1.6) and the variation found between

these. Judy's use of *stal, shtaul* and *stul* is found within the archive data and is found within sources that relate to self-ascribed speakers of Irish Traveller Pavee Cant (Dawson, 2011). The variant *stawl* is also found within the Manchester Corpus (2021) and attested as an Anglo-Romani item with cognate English *stall* 'stall'. *Starl* is also found in the Dawson Anglo-Romani word lists collected in Yorkshire and Arran (1959-69). This provided a suggestion of interdialectal influence. These two dialects remain distinct on a number of levels, however. In relation to Clinton's use of *ochi* and *akai* in *ochika, Stawl* is compounded with *akai* in the Manchester Corpus for 'stop', suggesting Clinton's use of *ochika* is not unique.

Variants for 'talk' elicited a uniform response with *rokker* and suggested a wider standardisation for this term. The term *vraker* means 'speak' in European Romani and is potentially a high frequency lexeme. Excerpt 6.4 illustrates the context of use for *rokker*.

Excerpt 6.4

1 Duncan: *as in talk is **rokker** and there's a load of people who wouldn't **rokker** to the none*

2 *[those outside the community]*

In this excerpt, Duncan illustrated the use of the variant *rokker* as integral to the status of being part of a community. This was expanded and contextualised further in excerpt 6.5

Excerpt 6.5

1 Duncan *they wouldn't speak because they just, you know, like you would get in the settled*

2 *community they wouldn't talk to a dead Gypsy, so well, that's not nice to say,*

3 *so I would say, you know to that person, I would say that's not a nice person*

4 *amongst their own people. I would say there's not a lot of people like that but*

5 *they, there is, there's people amongst our people that's not nice people.*

6 *There's good and bad amongst every people my friend*

This discussion related to the use of dialect and attitudes and beliefs within and outside communities. Duncan positions two groups as '*settled*' and '*our people'*. He used a metaphorical suggestion that '*in the settled community they wouldn't talk to a dead Gypsy'* an emphatic high-affective metaphor, and that this is '*not a nice person amongst their own'* and comparatively stating '*there's people amongst our people that's not nice people'*. This created

an appraisal of this potential tension as reconciliatory whilst acknowledging an awareness of prejudices that do exist that are outlined within Duncan's narrative.

Ben also used *puka* and this is also a variant found within the archive sources. *Pukker* is *phuk(er)-* 'tell' in European Romani and distinct from the English non-standard use of 'pukka' meaning 'good' related to the Hindi work 'pakka' meaning 'ripe' or 'mature'. Use of *puka* and *rokker* is plausibly interchangeable for 'talk' and 'tell' in the case of Jake. In Excerpt 6.6 below, the negator *kekka* was used alongside *rokker* and is an example of the negative marker *kek* or *kekka*. This is attested as the most frequent and productive grammatical marker in the Matras corpus (2010).

Excerpt 6.6

1 Clinton: ***kekka rokker*** *that means stop talking*

The form *kek* is associated with the cognate *kaj-jeck* from Sinti and Scandinavian Romani. Further details of dialect branches are discussed in section 2.2. and 2.4.7. It appears across a number of Anglo-Romani archive sources. The form *kek* and *kekka* both appear in the interview data and *kekka* is exemplified in excerpt 6.6 and further discussed in section 6.3.4 in the SRN category 'Phrases'.

Variables 'do/make', 'get', 'curse', 'tell' and 'run' elicited variants from three out of the six participants for each of these variables. Variants for 'do' and 'work' are comparable between Jake and Ben who produce '*kel*' and '*ker*' and appear across a number of Anglo-Romani word lists. The word *ker* means 'make/do' in European Romani and aligns with that reported by Jake and Ben. Clinton produced *buttei* and is shown in dialogue in excerpt 6.7. *Buti* is reported as 'work' in European Romani (Manchester Project, 2021). Although elicited as meaning 'make' or 'do', Clinton expanded his definition shown seen in excerpt 6.7. This example shows how lexis can refer to numerous concepts and also the possibilities for innovative usage within community and by individuals (Matras, 2010: 106; Elšík and Beníšek 2020: 416).

Excerpt 6.7

1 Clinton: ***butte*** *at it* ***butte*** *means work as well as something what you're working at…*

The variant is elicited in relation to variable 'do/make' and Clinton expanded on the meaning in relation to its use as 'work'. *Kel* may be a word within Clinton's lexicon although not elicited when prompted by discussion around the SRN.

‘Get’ is again elicited from Clinton, Jake, and Ben. These variants are dissimilar across participants. Clinton produces the phrase *latchet for nickcess* which did appear across a number of archive sources and cognates with the Cant term *niks(is)* meaning 'no’ in English (Manchester Project, 2021). This influence of Pavee Cant is characteristic of Clinton’s repertoire from the data, as well as considered a wider influence on Anglo-Romani (Matras, 2010). It also appears in Matras (2010: 140) as *rokker niksis* meaning ‘say nothing’. The word *latch* is a cognate of European *(a)(l)ač(h)/kh-* and means find in English. *Kin* represents a single attestation for the variable ‘get’. Ben use of *lela* /leilə/ is found as a variant within the archive vocabulary lists although Ben has an additional final vowel realisation proceeding lateral approximant /l/. *Lel* is attested as meaning 'take' in European Romani. The variant *astav* is not found amongst Anglo-Romani word lists within the archive and represents a variant with influence from an alternate Romani dialect. Except 6.8 below shows Clinton contextualising the use of the phrase used for variable ‘get’ in the context of ‘get for free’. This excerpt is taken from a wider discussion on lexis related to verbs associated with Anglo-Romani used by Clinton. Here, Clinton made a reference to spelling and alluded to the non-standardisation of Anglo-Romani spelling system as he states ‘*spell that’* in reference to a previous proposition from the interviewer to decide on orthographic representations. This epistemological claim positioned spelling as having a contentious status. Orthographic representation is humorously derided by Clinton as he challenged the interlocuters to spell this particular word. This suggested orthography as something that is symbolic to his identity as he is challenging a standard spelling system as part of this identity.

Excerpt 6.8

1 Clinton: ***latchet*** *for* ***nickcess*** *spell that,* ***nickcess*** *means nothing yeah*

The variants for ‘curse’ is not commonly found across word lists and this explains the variation between participants’ responses. Proximal or similar *Chingar* does appear in Dawson’s Derbyshire list as ‘curse’ and this aligns with Jake’s use, although more commonly the variant *chinger* aligns with English ‘upset’, ‘quarrel’, ‘swear’ and ‘insult’. The European *čhingerel* means 'quarrel (Manchester Corpus, 2021). Ben responded with *Waffado Lav*, and *Shavani Pukaras*. These are terms found in the archive as *Waffado* and translates as ‘bad’ (Dawson, 2002) and is reported as old Welsh Romani (Sampson, 1926a), *lav* meaning promise (Dawson, 2002) *alav* meaning ‘word’ in European Romani and so the term *Waffado lav* literally translates as bad/evil promise. *Pukaras* or *puker* means ‘to tell’ (Dawson, 2002) and the word *shavani* is

based on the Hebrew word *shav* meaning ‘return’ (Manchester Corpus, 2021) so the term *shavani pukaras* translates into ‘return telling’. These phrases do not appear within the archive lists as compounds and are original attestations. Clinton used the variant *a mock* although this does not appear within the archive material.

The variable ‘tell’ elicited variants *pucker*/*puka* (*puka* also discussed in relation to excerpt 6.5) and *pen* and are commonly found across Anglo-Romani word lists both historic and contemporary. The word *phuk(er)* means 'tell' in European Romani and is a cognate of these Anglo-Romani variants. The Anglo-Romani *pen* is *phen* in European Romani and translates as 'say'. The use of these terms are evidently interchangeable within Anglo-Romani usage as both pen and puka are reported for ‘tell’. ‘Say’ was not a variable and so it cannot be concluded whether Jake would differentiate between *puka* and *pen* without further elicitation. Clinton uses *boore* and gives context as illustrated in excerpt 6.9.

Excerpt 6.9

1 Clinton: ***boore*** *that means tell a lie talking yeah*

In Excerpt 6.9 Clinton explained the use of the variant boore as ‘tell(ing) a lie’ related but not synonymous with ‘talking’ which he also mentions. The use of *boore* for ‘tell’ is restricted to lie(ing) and again this example suggests a regional feature of Clinton’s repertoire as this term is not attested elsewhere although attested by Clinton as Anglo-Romani. It may relate to the European word for ‘big’ or ‘great’ being *boro/-I* associating with the meaning ‘exaggeration’ or extension of the truth and in this sense is an idiomatic extension of the word *boro/-*. Again, this idiomatic extension is a characteristic of Anglo-Romani which appears specific to certain communities and localities (Matras, 2010, Matras, and Tenser, 2020).

‘Run’ also elicited variants from three participants. Duncan and Jake used *gel/jell* which has a cognate in *džal(a)* in European Romani and means 'go’. This is also elicited as ‘go’ as described earlier. Jake also uses *nash* and relates to *naš*, meaning 'run (away)’. *Prasta* used by Ben also means ‘run’ and relates to the standard *prast* in European Romani. The variant *missly* was explained by Duncan as significant to growing up with two cultures and the significance of there being two distinct languages. Excerpt 6.1.1 shows Duncan explaining the variation within his own dialect repertoire and demonstrates his own awareness for distinctions he ascribes between the dialects of Anglo-Romani and Pavee Cant and how these indicate an association or characteristic of each community.

Excerpt 6.1.1

1	Duncan:	*that's what I was told there's lots of words like uh **missly** lots and lots of*
2		*different words*
3	Interviewer:	*missly*
4	Duncan:	***missly** again is to run or do you know and uh*
5	Interviewer:	*is that a word that you would have used growing up*
6	Duncan:	*yeah yeah because my mother's Irish and where we would say instead*
7		*of saying my father would say **luvver** which is for money my mother*
8		*would say **grade***

In terms of archive data, the variant *missly* does appear as *mislier* (Dawson, 2011) and is attested as 'walker', 'tramp' and 'nomad'. This is the only attestation for 'walk' and is not found in the archive lists amongst those used for this study but shows the association with the Irish Traveller community. In excerpt 6.1.1 Duncan also added his '*mother would say grade'* and showed the term *missly* would potentially be used differently depending on the context or concept.

'Know' elicited variants from two participants. Clinton produced *gen /ʤan/* with an open vowel /e/ compared to the Ben's closed vowel /i/ in *jin /* ʤ*in/*. The variant is shown in excerpt 6.1.2 Here, Clinton produced the variant in context and translates for the listener. This pronunciation aligns with one other variant from the archive sources and suggests this pronunciation may be a regional feature. Ben produced a more comparable word found in the archive data, with a closed high vowel in /ʤɪn/. There are a number of vowel variations within the archive sources, however. The variants *jin /ʤin/* and *gen* /ʤ*an/* correspond with *džin* meaning 'know' in European Romani (Manchester Corpus, 2021).

Excerpt 6.1.2

1	Clinton:	***dick ta ra gen…gen** no **gen gen** means I know or know you*

The variant *booty* and *buttei* were produced by Graham and Clinton respectively for 'work'. The example of use is given in excerpt 6.1.3 below, showing again the use of Anglo-Romani lexis within English morphosyntax, a characteristic of Anglo-Romani. The European Romani word for 'work' is similar being *buti* meaning 'work' (Manchester Project, 2021).

Excerpt 6.1.3

1 Graham: *you've got to **dick** a **mush** about some **booty***

The variant *yog* and *otch*er for variable ‘burn (something)’ were elicited from Clinton and Jake. The context of use is not specified in terms of any distinction between these two variants. Clinton did produce the term *yog* for the noun fire although is not used as a verb. In the example in excerpt 6.1.4 below, Clinton explaineed the use of the variant *yog* used as a noun and again shows the use of lexis embedded within English morphsyntax.

Excerpt 6.1.4

1 Clinton: *get that **yod** going get that **yod** going, what does that mean, get that fire going*

2 *__yod__ means fire what you'd do so when you'd put down a lane or build anywhere*

3 *and stopped. What the thing you'd do is get a fire going because you'd stopped*

4 *any animals from coming round wild animals because there's food and that*

5 *what you've got but you'll have food and smells in the wagon and one thing and*

6 *another so once they get a fire going you have a fire they won't come near you*

7 *you know what I mean. You get a fox or whatever like that they wouldn't come*

8 *near it once they get the smoke or whiff or something they'll stop away*

9 *they won't come to you*

In excerpt 6.1.4 above Clinton explained the significance of fire as a tool for itinerant travelling due to the natural circumstances of attracting wildlife. This references a cultural association of living in a customary peripatetic way. This association is built around the use of the term *yod,* and how it is used as well as the significance of this use. In this excerpt the discussion topic that stems from the word *yod* is related to cultural heritage. In the example, a connection between outdoor ‘*life stopped any animals from coming round wild animals’* and ‘*You get a fox or whatever like that they wouldn't come near’* builds an epistemological position in terms of word use and the culture that relates to these for Clinton. The variants *yog* and *otcha* are both found in the archive vocabulary lists. *Yog* has a cognate in the European Romani *jag* meaning 'fire' and *otcha* is close to *xač(ar)* which means 'burn' in European Romani. In the excerpt the variant *yog* is used as a noun and not a verb. Jake suggested this variant can is also

be found used as a verb. The cultural significance of the use of these variants is shown in the explanation tied to the practices associated with lifestyle and outdoor living.

Variants for ‘urinate’, ‘come’, ‘give’, ‘take’, ‘stay’, ‘raid’, ‘sell’, ‘to sleep’, and ‘read fortune’, were recorded by those ascribing to a Romani and an Irish Traveller dialect and were single productions. As discussed in section 3.5 of the methodology section, some of the variants were elicited through conversation and therefore it is not assumed these would not necessarily be part of other participants lexical repertoire. Following these words in Table 6.0, *mutter* features as a variant for ‘urinate’ and is produced by a single participant. In excerpt 6.1.5, Graham explained the use of an idiomatic expression in Anglo-Romani that relates to the use of *mutter*.

Excerpt 6.1.5

1 Graham: *so you could say if I were in a pub or I'd say I'm **geling** to the **muttering kenner***

2 *or **parnying** the **grai** it means I'm going to the piss house…water the horse so*

3 ***parny** is water which is like an Indian word isn't it*

Graham explained the term ‘*parnying the grai’* meaning ‘water the horse’ and then continued by explaining the connection between this and Indian language. This example demonstrates the use of Anglo-Romani lexis embedded within English morphosyntactic as the English suffix ‘ing’ is added onto *parny* in ‘parny-ing’ (Muyskin and Milroy, 1995). This is example of a characteristically inventive use of idiomatic and metaphorical way of communicating the action of ‘urinating’ as ‘giving a horse some water’. It again also showed an awareness of the heritage of the dialect as historically originating from the sub-Indian continent by speakers of the dialect as Graham mentions ‘*parny* is water which is like an Indian word isn't it’. The variant *mutter* is found in a number of archive sources and translates as *mut(e)r-* in European Romani and means 'urinate'. Another single production is the variant *crush* used by Duncan. In excerpt 6.1.6 below, he explains the significance of the use of the variant as relating to his own identity, being from a mixed heritage background outlined in section 3.3.1. The word *crush* was not found within Anglo-Romani vocabulary lists as expected for speakers of Anglo-Romani dialect and is instead found in the Iris Traveller Pavee Cant vocabulary list (Dawson, 2011). This variation relates to his self-ascribed mixed heritage identity and in excerpt 6.1.6 was positioning himself as identifying with both cultural associations (Bucholtz and Hall, 2005).

Excerpt 6.1.6

1 Duncan: *yeah there's a lot of words we'd use. Like I did tell you my mother's Irish so*

2 there's a lot of words that we'd use. Like you know for speaking and for ***crush***

3 instead of the word ***gel*** *which isn't English the* ***Romanesk*** *word they would say*

4 ***crush…***

5 well we'd say uh ***crush*** *a if it was a boy we'd say or you were speaking to a man*

6 ***crush*** *here* ***sooble*** *so it would be come here boy or come here man you*

7 know…so that would be Cant Irish Cant

In excerpt 6.1.6 Duncan explained how *crush* meaning 'go' is used (*gel* is discussed in Chapter 4, excerpt 4.1), suggesting Duncan's identity is constructed intersubjectively as he positions his background as mixed-heritage (Bucholtz and Hall, 2005). Duncan described his heritage discussing Irish Traveller and Romani as part of a cultural heritage. In the discussion he explored the use of words that he uses: '*for* ***crush*** *instead of the word* ***gel*** *which isn't English the* ***Romanesk****'*. He continued with the example '*crush here sooble'* meaning 'come here boy', depicting his background and a cultural tie with both identities.

'Give' is produced as *del* by Graham and appears across a range of historical and contemporary sources referenced in Table 6.0. This again suggested a degree of maintenance of this variant although is not elicited from the other participants. There is an expectation that high frequency words such as 'give' would be a predictor of higher maintenance, but this is not evidenced here. Instead, the notion that other more significant factors that relate to identity and practices of identity are born out with the degree of variation in attestations and also difference in variants. *Del* is 'give' in European Romani as is *d-* (3SG) and is proximal in terms of orthography comparing the variant used by Graham. In excerpt 6.1.7 Graham was explaining the term *luvver,* which is discussed in detail in section 5.1.3, and uses the verb *del* to give context to the use of the variant for 'money'. Graham in this excerpt appeared to distance himself from the use of lexis by using the unreal conditional aspect in *'you'd'* positioning himself as not central to the use of the associative word. Whilst discussing his cultural ascription and ties to Roma culture he also positioned himself as an observer within the interview. As stated, this may be contextually elicited in relation to the formal context of the interview or a nuanced feature of Graham's Anglo-Romani identity.

Excerpt 6.1.7

1 Graham *its more likely you'd hear **luvver dell** the **luvver** give the money*

'Take' is attested as *puv* by Duncan and is not found in the Anglo-Romani word lists within the archive sources. In excerpt 6.1.8 Duncan explains the use of the variant *puv* and demonstrates Duncan adding to the SRN's 'and other' category.

Excerpt 6.1.8

1 Duncan: *to take something, I don't know how you would say you haven't got, take there*

2 *have you the word take as I take is again it's uh **puv** to take something is*

3 ***puv…**uh again this one to go or to run or away would be uhm to stay in a place*

4 *can I say sit here I would say **atchie artch** means to stay here **atch atch** here*

5 *and for example years ago when people used you stop down thier wagons they*

6 *would say this is an **atchin tan***

With reference to the archive data sets, the variant *puv* is not found in the Irish Pavee Cant vocabulary lists within the archive or the Leeds recordings. The nearest approximation is *bog* meaning 'take' in Irish Traveller Pavee Cant (Dawson, 2009) with this word for 'take' an original attestation. Although *puv* is not listed for 'take' in the archive sources, a number of similar variants are listed within the archive sources for verbs 'make' and 'do', having a number of alternate variants in Irish Cant and Anglo-Romani. In excerpt 6.1,8, Duncan continued to explain the variant *atch* which is recorded as *ač*- meaning 'stay' in European Romani (Manchester Corpus, 2021). Graham, Clinton, Jake and Ben attested the use of *atch* as used for 'stop over' and 'camp'. The use of *atch* therefore varies in relation to the context of use as Duncan used *atch* for 'stay', whilst the use for *atch* to mean 'camp' or 'stop over' is also attested with the data. This demonstrated the degree to which meaning for lexis for Anglo-Romani varies between individuals and the extent to which one variant can refer to a distinct concept.

A further single elicitation and a variant not recorded as an Anglo-Romani or Cant is Graham's use of the term *steaming*. In excerpt 6.1.9 the term is contextualised in relation to events in a Gypsy fair and is comparable to some of the terms Judy uses that associate with vernacular English.

Excerpt 6.1.9

1 Graham: *ah because she had some problems up there didn't she with what they call*

2 ***steaming****…*

3 *where this was Irish Traveller lads that charge into wherever you are and grab*

4 *anything they can and off they go and they call it steaming*

From excerpt 6.1.9 and with evidence from the archive sources, this seems likely a recent phenomenon and coinage. A further single elicitation from Graham is for the variant for 'hawking' and is given as '*monging'.* In this excerpt Graham explains in some detail the type of selling this refers to.

Excerpt 6.1.9 continued

1 Graham: *well some for the older words like to go out hawking to go out hawking and as*

2 *in you'll still see stickers doors saying no hawkers that's that means really for*

3 *door knocking is selling isn't it. I'm going out hawking I'm going out calling*

4 *would be like a scrap mans work I'm going out calling for scrap like any old*

5 *iron…calling as in like you would going round in the vans looking for scrap*

6 *shouting any old iron that sort of calling isn't [looking for old iron] it's looking*

7 *for tree work or…****monging*** *yeah*

In excerpt 6.1.9 Graham referred to 'hawking' as a potential lexical variant used by Anglo-Romani speakers in *'I'm going out hawking I'm going out calling'*. Graham also associated the word '*monging'* with doorstep trading 'to hawk'. The variant *monging* was represented in the archive proximally as *bikomengro* although this is less similar than that found in the Manchester Corpus and is listed as *mongipen.* The European Romani *mangipen* means 'request' in English and is a cognate for this variant. The word for 'sleep' was elicited from one participant and was found across a number of archive sources. The variant aligns with the European Romani *suto* meaning 'sleeping' or 'asleep'. The variant is explained in the context given in excerpt 6.2.0 below. The variant is embedded within English morphosyntax with use of the suffix 'y' as in 'gone to ***sutty***' and where the dialect variant is embedded within the infinitive verb as in 'gone to'. The embedding of lexis within morpho-syntax is commonly utilised in codeswitching (Myers-Scotton, 1993) and would be expected for speakers using

both Romani and English and hence the predominance of Anglo-Romani having this characteristic.

Excerpt 6.2.0

1 Duncan *gone to **sutty** in English we would say you have gone to **sutty** gone to sleep if if*

2 *where'se Michael oh he's gone to **sutty** or where's me wife she's gone to **sutty***

3 *she's gone to sleep okay*

Dukkering is explained as fortune telling and is discussed in the context of an extended dialogue on Gypsy fairs and events and practices that take place. Excerpt 6.2.1 highlights Graham explaining the perspective of the fortune teller and the use of the term in relation to use of the variant *ducker* meaning 'to tell a fortune'

Excerpt 6.2.1

1 Graham: ***duckering** that's like fortune telling…*

2 Interviewer: *oh **duckering** I've really not heard any of these is that something that*

3 *you come across*

4 Graham: *you will if you go to fairs or anything like that you'll come across alot of*

5 *fortune tellers but amongst themselves they would [say] you gonna*

6 ***ducker** somebody or make some…*

This epistemological proposition relates Anglo-Romani lexical variant use with covert communication (see discussion on excerpt 6.2 and 6.3.6). Graham used the third person reflexive pronoun referent '*amongst themselves'* which showed a nuanced relation in terms of his affective stance towards use of dialect and his own use.

6.1.2 States

The second most productive category in the SRN 'Feelings, Actions and States' is 'States'. 'States' refers to adjectives or descriptive terms. Table 6.3 shows a total of 33 variants for a total of 8 standard English variables. Variables are elicited from the SRNs with associated variables and variants raised through discussion displayed either alongside these variables. Where there are distinctions in terms of semantic meaning, these are displayed in a separate row. As with all SRNs the proceeding Table (6.3) represents the code for the references

displayed underneath the variants within Table 6.3. The complete bibliographic references for these can be found in the reference list.

Table 6.3: SNS: Feelings, Actions and States: **States**

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Good			*Latchi/Baari* MC, DV2/DV1, DV3	*Kushti bokka*	*Kushti* DV2, EA, MC, DV1,DV3,D1,LSP, V1	*Kushti/latcho* DV2, EA, MC, DV1,DV3,D1,LSP, V1 /W3, MC
Lovely			*Latchi/Baari* MC, DV2/DV1, DV3			*Rikani/shuka*
Ashamed			*Ladge* SC1, D1, S1, W1	*Ladged* MC, SC1	*Lajd* MC, SC1	*Alardge/Lajilo* SC1, S2/
Stolen/Steal			*/Chorey* SC1	*Chored/* MC	*Chored/Chore* MC/DV2,T1,V1,B2	*Chorlo/Chor* / DV2,T1,V1,B2
Mad/crazy				*Radged* MC,DV3	*Dinilos/Divo* IV/D1,S1,W1,B2, SC1	*Ragged/choro mush* MC,DV3/
Big		*Barri* DV2	*Boot* R,G,EA, S1,	*Bore* DV2	*Bawri/Bori* DV2/ DV2	*Baro* MC, V1, B2

Table 6.4: Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

The variable that was most productive in terms of variants produced was 'big'. Six out of the seven participants produced variants for this variable. The variants *barrey /bæːɹi/* and *bawri /baːɹi/* produced by Graham and Clinton are comparable but distinct in terms of the initial vowels /*aː*/ and /*æː*/ represents a proximal phonological variant (see section 3.5 for discussion on variation analysis and treatment of spelling). *Bore //baːɹe/*, *baro /bæɹaʊ/* and *bori //baɹi/*, are also similar but with distinctions between initial and final vowels realisations. All of these variants are represented in the archive data, although *bore* is limited to the Derbyshire vocabulary list (Dawson, 2002). There is a suggestion of regionality at play as Clinton described himself as a long term resident and from the Nottinghamshire area. This could therefore represent a regional characteristic in the realisation of the variant *bore*. Jake noted that the variant can be realised as *bawri /baːɹi/* and *bori /baɹi/*, representing two distinct orthographic and therefore representing alternate phonological realisations. This indicated acknowledgement of variation existing in orthographic realisations representation by Jake, in addition to there being phonological distinctions. However, no standard orthography has been coded for Anglo-Romani, although work is being done to form a standard orthography and a standardised Romani (Matras and Leggio, 2017). These five variants have a cognate with the European Romani term *baro/-i* meaning 'big or 'great' in English. Excerpt 6.2.3 shows how it is used with *bari*, used as an adjective to describe a 'big man'. *Barri mush* in this excerpt, is an

example of lexis embedded with English syntax as in '***dick** the **barri mush**'*. The grammatical frame follows the English construction for the imperative. The commanding verb is followed by the object with notable exception in the dropping of the preposition 'at'. This phrasal distinction between English and the non-standard variation is demonstrated in the example below. Graham indicates an understanding of grammatical contrasts between the use of Anglo-Romani with the alternating pattern of use or codeswitching practices of the mixed language of Anglo-Romani.

Excerpt 6.2.3

1 Graham: ***barri mush** like **barri mush** look at the big man, **dick** the **barri mush***

The variant produced by Duncan, although used for 'big' (*boot /bu:t/*) is distinct in terms of vowel quality and final consonant. The variant *boot* although used for 'big' as shown in excerpt 6.2.4 cognates with European Romani *but* meaning 'much' or 'many'. The use of *boot* as an adjective to modify countable and uncountable nouns is evident in its use in excerpt 6.2.4 in '***boot luvver'*** and shows evidence in the data of the distinction between *bore* for 'big' and *but* for 'much' and is evidenced here as maintained in Anglo-Romani in modern usage.

Excerpt 6.2.4

1 Duncan: *that word there would be **boot** it's Romanes **boot** means big .., yeah well we say*

2 *money I told you the other day money was **luvver** yeah okay so if it was big*

3 *money we would say **boot luvver***

Three variables produced variants from four of the participants. These were 'good', 'ashamed', and 'stolen'. The variants for 'good' given by Duncan and Ben, *latchi* and *latcho* both appear in the archive vocabulary lists with several variants appearing across archive sources with proximal orthographic realisations.

In excerpt 6.2.5, Duncan explained the use of both *baari* and *latchi* are interchangeable. He continued to speculate that these two terms originate from two separate sources, indicating how these two distinct variants have come to coexist. This discussion showed engagement with the heritage of the dialect variants and his self-ascribed association with Romani identity and background outlined in section 3.3.1. This related to the relational analogue scale (see Table 4.3) in terms of his self-ascribed identity which is complex, composing his relational identity. His Anglo-Romani identity is realised through a meta-linguistic narrative and is part intentional,

whilst constructing an identity through a narrative constructed through his experience within the frame of the interview (Bucholtz and Hall, 2005: 606). In this sense, Duncan was building identity through negotiation and contestation of the historical accounts of these variants.

The term *lačho/-i* meaning 'good' in European Romani and *baro/-i* for 'big' or 'great' are cognates and etymological sources for these two terms.

Excerpt 6.2.5

1 Duncan: *okay no problem **baari** good is **baari** or **latchi** another word is **latchi** there's*

2 *two different words for the same thing yeah sometimes we have… I think what's*

3 *happened here with these different here is eh again I think they've been different*

4 *words from different countries like **latchi** and **baari***

The principle that Anglo-Romani words are attested from different sources is found within the literature as words from Welsh-Romani, Sinti as well as more ancient influences from Greek and Iranian (discussed in section 2.4.7). Ben also produced the variant *latcho* and is proximal to Duncan's *latchi.* Both align with a number of proximal variants in the archive lists, as well as matching variants within the archive lists. The variant *kushti* was elicited from participants Duncan, Jake and Ben and all three are matching in orthographic realisation across several archive sources.

In excerpt 6.2.6 Clinton gave context for the use of the adjective *kushti* used together with variant *bokker* as in *kushti bokker.* The European Romani *kuč* means 'expensive' and the variant *kushti* is found across several lexical lists and commonly used as an adjective meaning 'good'. This is an example of word that has spread outside of the Anglo-Romani community of use although maintained as an in-group lexicon. The word *bokker* was not found amongst the archive lists or corpus collocated with *kushti* and represents a unique attestation in terms of collocation with the variant *kushti,* (see section 7.3.3 for *kushti bokker*). The variant *bok* is found across several sources (Dawson, 2002) and is attested as corresponding with English 'luck' and 'fortune' and corresponds with European Romani *baxt* and equates with English 'luck'. The most proximal attestation is *bawker* (Dawson, 2002).

Excerpt 6.2.6

1 Clinton ***kushti bokker** means good… **kushty** no not **kushty** oh fucking hell its gone out*

2 *me head… **dick ta ra gen… gen** no **gen gen** means I know or know you*

The English variable ‘ashamed’ elicited six variants from four of the participants. The variants *ladged* and *lajd* were used by participants Clinton and Jake respectively. These were similar or proximal in their realisation. Near cognates are *ladgin*, and *alardge*, are proximal, and both appear in the archive vocabulary lists closely relating to those used by Clinton and Jake. *Lajilo* given by Ben however, does not appear in the archive Anglo-Romani word lists and represents a variant found outside of Anglo-Romani dialects. This is evident from its omission from the archive and corpus data. This aligns with several variants used by Ben which do not appear in the Anglo-Romani archive lists. The cognate term for all these variants in European Romani is *ladž-* and corresponds with 'be ashamed' in English as well as *ladžardo/-ano/-alo*; also 'ashamed'. In excerpt 6.2.7, Duncan explained the use of *ladge* meaning ashamed amongst descriptions of variants for ‘alcohol’ and ‘good’. He explained the possibility of sources of origin and how both variants *barrie* and *latchi* could be used interchangeably.

Excerpt 6.2.7

1 Duncan: *and it's again it's they told me it's polish so that's what I've been told this word*

2 *here shame means **ladge ladged** when your **ladged** when you're ashamed I'm*

3 ***ladged** I'm ashamed yeah uh uh let me say I would put this on the same word as*

4 *this uh good again uh lovely **latchie barrie barrie latchi** actually it's the same*

5 *word **barrie***

In excerpt 6.2.7 Duncan moved between the claim of a term being a Polish in word ‘*it's they told me it's polish so that's what I've been told’* which acts to distance himself from the claim with modality evident in relation to this claim of origin. He continues by focusing on the distinction between the two terms in ‘*latchie barrie barrie latchi actually it's the same word barrie*’ and authenticates this proposition with the declarative ‘it’s’ as Duncan authenticates the proposition.

‘Stolen’ or ‘steal’ elicited variants *chorde /ʧa:d/* and *chored /ʧa:d/* from Clinton and Jake. These are both proximal in spelling and phonological realisation. Given the non-standard orthography this might be expected. *Chor* and *chore* are proximal realisation used by Jake and Ben for ‘steal’, while *chored* and *chorlo* were used by Clinton, and Jake and Ben for ‘stolen’ respectively. *Chorey* was used for ‘stolen’ by Duncan. From the archive sources for Anglo-

Romani all of the variants found amongst the participants are listed except for Ben's use of *chorlo*. This again is proposed as a non-Anglo-Romani variant and attains to another branch of Romani dialect. As discussed, this is a characteristic pattern for Ben's variety. The European Romani word *čor*-is a cognate meaning 'steal' and also *čordino* meaning 'stolen'. The variants *chored* shows English morphological '-ed' ending embedded within the Anglo-Romani variant, whilst *chorey* does not. Excerpt 6.2.8 illustrates Duncan's explanation of the use of *chorey* as meaning stole. The *-ey* ending is a unique attestation and represents an individual coining or regional variation.

Excerpt 6.2.8

1 Duncan: *let me think here now let me have a look here now uhm let me think*

2 ***chorey*** *in this one stolen* ***chorey*** *okay*

In excerpt 6.2.8, Duncan considers the meaning of the word *chorey* in '*let me think here now let me have a look here now uhm'* before declaring '*this one stolen chorey,* suggesting lexical variation is not immediately recallable for some less frequently used words.

Table 6.3 also showed the variables 'mad' and 'crazy' produced variants from three of the participants. The variants *radged /ɹædʒid/* and *ragged /ɹædʒid/* are proximal from Duncan and Ben and appeared in the archive and corpus phonetically transcribed as ['ɹædʒi]. However, there is no European Romani attested. This variant is specific to Anglo-Romani as there is no clear Romani etymology for this variant and shows Anglo-Romani is not only a language that borrows from European branches. The variants *dinilos* and *divo* given by Jake are also found across Anglo-Romani lists and correspond with European *dinilo* meaning 'fool' and *divjo* meaning 'wild'. The English vernacular 'divvy' is translated as 'stupid' and has an etymological connection with the European Romani variant *divjo* and Anglo-Romani *divvi* (Matras, 2010).

A variant for 'lovely' was produced by two participants. Ben produced the variants *rikani* and *shuka* which were elicited both for the variable 'lovely'. Neither of these variants appear in the archive lists and are hypothesised as two variants representing an alternate branch of European Romani and not Anglo-Romani. Duncan produces *latchi* and *baari* commenting that these are used in the same way to mean 'good' although elicited with the variable 'lovely' illustrated in the previous excerpt 6.2.7.

6.1.3 Feelings

The third most productive category within the SRN 'Feelings, Actions and States' was 'feelings' in terms of the number of variants produced across participants and displayed in Table 6.5. There were a total of 18 variants elicited from a total of 6 standard English variables which were raised through discussion around the SRN. Table 6.5 below displays the variables and variants. Variables correspond with variants in relation to their position either side of the forward slash. As with previous Tables of results, the letters represent approximate matches with archive sources with the Table of codes given below in Table 6.6.

Table 6.5: SNS: Feelings, Actions and States: **Feelings**

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Tired		*Paggered* IV, DV2,		*Suttery* EA	*Kinlo* S1, V2, SC1, MC, D1	*nai kushti* MC, SC1, W3, R
Pleased/Happy				*Kushty bot* MC	*Kushti/Bahlali* DV2, EA, MC, DV1,DV3,D1,LSP, V1/DV3	*Kamiyado*
Sick				*Gammy* MC	*Nafelo* S1, SC1, N1, DV5	
Angry	*Crabbed*			*Radged* MC,DV3	*Rusdo* MC, DV3	
Scared				*Shadder*	*Atrasht* MC, DV1, D1, F1, S1, B2, SC1,S2, DV5	*Trashin* MC, DV1, D1, F1, S1, B2, SC1,S2, DV5
Drunk		*Peeved* MC			*Muti* *DV3*	

Table 6.6: Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

The variable 'tired' elicited a response from four of the participants. All four of these variants are different in terms of etymology, indicated with reference to the archive sources. Graham produced the term *paggered* which is found in the Derbyshire word list (Dawson, 2002). Variations on this variant are found across a number of Anglo-Romani sources such as *pogger* (Manchester Corpus, 2021) and *poggerin* found in the Leeds Special Collection recordings (1979). However, the realisation *pagger* is limited to word lists of the East Midlands. The European Romani word phag(er) is a cognate and means 'break'. This regional characteristic again is suggestive of variation influenced by locality of speaker and conclusions drawn from this are summarised in section 8.2.1. The word also has a number of meanings associated with old Welsh Romani such as 'to train' or 'tame' an animal (Manchester Corpus, 2021; Sampson, 1891) and again this shows a pattern of extension of lexical items within Anglo-Romani. In excerpt 6.2.9 Graham explained how the use of *paggered* for 'tired' is understood widely within the dialect speaking community and gives alternate meaning, using 'fed up', 'down', and 'beat up'.

Excerpt 6.2.9

1 Graham *or if you were really really fed up or down you'd you could say I'm* ***paggered***

2 *couldn't you…I'm **paggered** I'm beat up you know…they'd all know what it*

3 *meant if you said I'm **paggered** they's all I don't know any that wouldn't know*

4 *what you meant they'd know exactly what you meant*

Duncan was able to produce associative terms through both onomasiological (what words associate with a specific meaning) and semasiological (what one word can mean) demonstrating a finer grained analysis of lexical variation often problematised within the study of lexical variation (Durkin, 2012). Although the realisation *pagger* is not common across Anglo-Romani word lists, evidence shows this word is widely recognised across communities of Anglo-Romani speakers (Manchester Project, 2021). In this example, Duncan makes a positional statement in '*I don't know any that wouldn't know what you meant'* determining the use of this lexis with membership of a group. The proposition expresses a degree of epistemological certainty, positioning the term common amongst a community he ascribes as part of his own identity (see fig. 4.3).

Table 6.4 showed Clinton produces the variant *suttery.* This appears as *sutti* in the East Anglian Anglo-Romani word list and appears across a number of Anglo-Romani word lists although not used as an adjective but as a verb. The word *sutti* is a cognate of European Romani *suto* meaning 'sleeping' or 'asleep'. In excerpt 6.3.0, Clinton explained how the word is used in context, with English morphosyntactic adjective ending '-ery' used to derive the word *sutti* as an adjective.

Excerpt 6.3.0

1 Interviewer: *I hope you're not getting tired feelings tired when you're knackered*

2 Clinton: ***suttery** sort of means sleep*

This use of English morphosyntax is a feature of Anglo-Romani as highlighted and discussed in reference to Excerpt 6.2.0. and again highlighted through this example.

Jake was consistent in his productivity and in producing different variants from the other participants with *kinlo /kinlɑʊ/*. Similar variants are found across a number of sources although this example is the only occurrence of lateral approximant /l/ proceeding the alveolar nasal which is unexpected in relation to the archive data. This was a feature of Jake's repertoire that may relate to his background being from Lincolnshire and identifying more closely with 'British Romani' (see Table 4.1 and fig 4.3). This suggested a regional or localised feature for

production of the variant. The European cognate *khino* means 'tired' (Manchester Corpus, 2021). In contrast, Ben produced the variant *nai kushti*. The variant *kushti* used for ‘good’ is explored in section 6.1.2 and illustrated in excerpt 6.3.1. outlined as an example of an Anglo-Romani term that has spread outside the community. There was no suggestion of an attitudinal shift towards its use in the discussion excerpts. This suggested self-ascribed speakers of Anglo-Romani within this study still use the term as indexical of Anglo-Romani identity. The adverb *nai* was used to negate the word *kushti* is used twice amongst the corpus of this project and appears across a number of archive sources as a functional lexeme meaning ‘no’ and ‘not’ (Manchester Corpus, 2021). The variant is a cognate of the European Romani word *na* meaning 'no’ (Manchester Project, 2021). The variants elicited for ‘tired’ shows a degree of productivity for Anglo-Romani speakers. This variation characteristics is possibly a result of being an uncoded language, and also the results of eliciting lexical variation using a both onomasiological and semasiological approach to lexical elicitation.

The variables ‘pleased’, ‘angry’ and ‘scared’ all elicited variants from three of the participants. Clinton produced the variant *kushti bok* for ‘pleased’. Excerpt 6.3.1 illustrates the context of elicitation and the context of use explained by Clinton.

Excerpt 6.3.1

1	Interviewer:	*when you're happy proud pleased chuffed*
2	Clinton:	*you'd say* ***kushty bok*** *what if it were me who were feeling happy…* ***kushty***
3		***bok*** *means feeling happy*
4	Interviewer:	***kushty bok*** *sick under the weather nackered poorly*

This example illustrates meaning can be dependent on the context of use. The use of *kushti* also shows a more nuanced use for those that ascribe to Roma and Traveller identities or those that identify as a speaker of Roma or Traveller language or variety. The variant *kushty* is used across a number of contexts meaning ‘good’ as described in 6.1.2 and used again here as an adjective describing the state of ‘happy’. The collocation was found within the archive in *kushti bok* in the Manchester Corpus (2021) and cognates with European Romani *kuč* meaning 'expensive' and *baxt* meaning 'luck'. The associative cognate terms shows an indirect connection between there meaning and the use of *kushty bok* and an example of the characteristic of innovation for Anglo-Romani and more widely a contact feature of Romani (Tensor and Matras, 2020). In

this excerpt, Clinton said '*if it were me…*' which within the context can be positioned as indirectly ascribing an Anglo-Romani identity with the use of the phrase '*kushty bok*'.

Clinton explained the extended use of the lexical variant to describe the state of being 'happy' and 'pleased'. This expression of use is not attested within the archive Anglo-Romani word lists. The word *kushti* used for 'good' is found across a number of archive sources and again is found to have been extended in use by Jake. Again, this extended use for 'please' was not found within the archive lists. Jake also produced the variant *bahlali* which has a proximal and associative variant (see section 3.5) in *bahtale* found in the Dawson's 2009 word list which references words elicited by self-ascribed speakers of Scottish and Irish Traveller dialects. One conclusion from this archive source could be the variant stems from Irish or Scottish Traveller dialect etymology or has a cognate in the European Romani word *baro* meaning 'big' or 'great'. Ben produced the term *kamiyado* which is not found in the archive lists and represents a variant from outside the Anglo-Romani data sources. This could mean this is either a variant from a separate Romani branch (see section 2.5 for details of Romani dialect branches) or unique innovation to Ben and relative to his background (see section 4.1). 'Angry' elicited three variants and significantly the only variant from Judy in this subcategorisation from this SRN. Judy is predominantly less productive. In terms of background data, Judy's Identity Score Index was lower at 4/14 and one factor related to her productivity of non-standard words. The variant was found across a number of sources as relating to the English word 'angry' although not suggestive of either Irish Traveller or Anglo-Romani heritage. Clinton produced the variant *radged* indicating how this variant is used in various contexts and can be used to mean angry as well as for 'fool' or 'foolish' (see section 7.4.4 'Personality').

The variant *radged* appeared across two Anglo-Romani sources and translates as 'mad' (Manchester Corpus, 2021). Excerpt 6.3.2 illustrates Clinton's use of the variant in this context of use. The variant *radged* is not recorded as used for 'Angry' and so this is a unique attestation. Jake produced the variant *rushdo* and is found within the Manchester Corpus (2021) as *roshta* and also in the archive lists as *roshto* (Dawson, 2009). The Europen Romani Cognate *rušto* means 'angry' (Manchester Project, 2021). The current variation found within this project's corpus illustrates variation within Anglo-Romani, suggesting a nuance and complex picture of those that self-ascribe to either having Roma, or Traveller heritage as well as those that describe their own identity using self-appellations such 'Roma' and 'Irish Traveller'.

Excerpt 6.3.2

1 Interviewer: *angry or mad upset with something*

2 Clinton: *again its eh it'd be you know same as what I said before…*

3 ***radged*** *going* ***divvy*** *yeah*

In the above example, Clinton demonstrated again the multiple meanings of terms that can are given so 'angry', 'mad' and 'upset' are all related to the term *ragged* or *divvy*. The term ragged has also been adopted or borrowed into mainstream regional vernacular or 'slang', and not influenced patterns of use amongst Anglo-Romani as shown in the example above.

'Scared' also produced variants from three participants being *shadda, atrasht* and *trashin* from Clinton, Jake and Ben, respectively. There are several proximal variations across the archive Anglo-Romani word lists that align with Jake and Ben's productions of *atrasht* and *trashin* respectively. The European Romani cognate *traš* means 'frighten', meaning 'scared' in English. The elicitation of *shadder* by Clinton is not found with the archive material however and illustrated in excerpt 6.3.3. The variant *chad* is attested as 'chased away' as in 'I got *chad* last night' (Manchester Corpus, 2021) and is the nearest attestation aligning with *shadder* for 'scared'. No corresponding European Romani is attested to this variant, so representing a unique attestation for Clinton shown below.

Excerpt 6.3.3

1 Interviewer: *yeah scared or frightened afraid?*

2 Clinton: *uh* ***shadder***

3 Interviewer: ***shadder*** *so how would you say that*

4 Clinton *I don't like this I'm frightened of that do you know what I mean…*

5 *yeah I'm* ***shadder shadder***

The last two variables 'sick' and 'drunk' elicited variants from two participants, respectively. Clinton used the term *gammy* attested as English vernacular (Manchester Project, 2021). In excerpt 6.3.4 the dialogue relates to Clinton's explanation of the variant *gammy.* In the excerpt 6.3.4 the interviewer made a declarative statement regarding the status of the variant *gammy*. He positions the use of *gammy* as used by non-speakers of Anglo-Romani as he states 'we' positioning those outside the Roma community. There is a suggestion of 'othering' (Prieler, 2010) by the interviewer with use of '*we use gammy'*. Clinton overrides the definition proposed

by the interviewer for *gammy* and continues '*gammy gammy means poorly…'*, taking an authoritative stance in respect to the meaning of this variation.

Excerpt 6.3.4

1	Interviewer:	*any other feelings?*
2	Speaker 2:	*yeah* ***gammy*** *leg we use* ***gammy***
3	Clinton:	***gammy gammy*** *means poorly…* ***gammy*** *he says*
4		*oh I'm* ***gammy*** *I'm* ***gammy***

This term was not identified within the literature as Anglo-Romani and reflected here by Clinton who does not explicitly relate the word to Romani.

Jake used the variant *nafelo* which is found across a number of archive sources in proximal form, and 3 instances of near realisations. The European Romani *nasvalo* is a cognate and means 'sick' or 'ill'. 'Drunk' also elicited variants from 2 participants. 'Peeved' is used by Graham for 'drunk' and has a European Romani cognate *pijav(a)* which means 'drink'. The English morphological inflectional ending -ed changes the word meaning from 'drink to 'drunk' and reflects the innovation related to language contact for Anglo-Romani with Romani words embedded within English morphosyntax (see section 2.4.7).

In excerpt 6.3.5 Graham gave an extended narrative contextualising *peeved*. Graham's narrative suggested thematic topics related to morality, generosity and affluence are culturally integral. Respect for loyal and moral action is central to the story as the *'old Irish horse dealer'* is portrayed as a deserving action.

Excerpt 6.3.5

1	Graham:	*I mean we were, I were in a pub once drinking all day down in the town like al*
2		*l all day until they through us out of the pub because it were closing and we*
3		*were walking back up the hill which was at least about a mile, and a half two*
4		*miles isn't it and I were with some o' old Irish horse dealers and one of them*
5		*fell into somebody and knocked the drinks all over. So this bloke were going to*
6		*have a right go at him and I stood in the middle of them and I said look the*

mush's peeved *you know he don't mean any harm we're going to sort this thing alright alright he says I'll pay for his drinks he did and they went we walked back up the hill holding on to one of each arms walked up the hill got 'em back to their camper, and in the morning they were stopping. We were like on the hill, they'd called us over me my friend and he said come in we'll do you your breakfast come in…*

The deserving act being the action of stopping an altercation as Graham explained '*I stood in the middle of them and I said look the mush's peeved you know he don't mean any harm…'.* This action is then rewarded by the 'old Irish horse dealer' as Graham explains ' *I want to buy you something for last night as a thank you for that last nigh…'*. Generosity is arguably central to the story and to an affective attitude towards Roma identity as Graham quoted the man saying '*pick something on the fair what you want and anything I looked at whether it be a pair of boots or anything he said let me buy them for you…'* in line 17.

Excerpt 6.3.5 (continued)

So we were having breakfast in this big like winnibego camper and he did a sirloin steak breakfast and all this business. And he said to me we're going to have a walk round the fair then we're going back to the pub alright so he said I want to buy you something for last night as a thank you for that last night So pick something on the fair what you want and anything I looked at whether it be a pair of boots or anything he said let me buy them for you. I said I don't want anything, I don't want anything and he went like that in his pocket and he pulled out a wad of money like that and it was that deep folded up thousands and thousands and thousands and thousands and he said if it weren't for you that wouldn't be in my pocket this morning…So he said let me buy you something from the fair. I said I don't want anything I don't want anything. I said I'll tell you what go and buy me a can of coke from the icecream van and

25 *then we went to the pub and did it all again.*

This showed positive evaluation towards Roma culture. The notion of affluence is also a noticeable focus. Graham explained that the *'old Irish horse dealer'* in line 20 was pulling *'out a wad of money like that and it was that deep folded up thousands and thousands and thousands and thousands ….*'. Arguably this affective appraisal builds a positive account, positioning a cultural identity within the interaction.

Jake uses the variant *muti* and this is found in in proximal form (*moti*) in Dawson's list of words collects from Derbyshire Travellers (Dawson word lists; Special collection, 1989-91,). There are several less proximal variants and relate to the European Romani *mato* meaning 'drunk' in English.

6.1.4 Phrases

The final subcategory within the SRN 'Feelings, Actions and States' is 'phrases'. The phrases were chosen to elicit function words and is the only subcategory that purposefully identifies these. Whilst function words were dealt with ad hoc through the rest of the analysis (section 4.2.2) this section specifically looks at phrases and the use of function words.

Table 6.7: SNS: Feelings, Actions and States: **Phrases**

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
I have no shoes					*No Chitiaws Chokas*	
Shut up		*(Kekka) Chave* (R,G, EA, T1)		*(Kekka) Rokker* (R,G, EA, T1), MC	*Chichi* DV2, S1, W1, B2, SC1	*(nai) puka/ (ma) puka/ (kek) Rokka* (MC, W2, R)DV2, V1, B2/(R) DV2, V1, B2/(R,G, EA, T1), MC
Not		*Kekker/kek* R,G, EA, T1				

Are you a Roma/Traveller?				*A Rumnus gearer* (fill in from below)	*Pukker your nav* (MC, DV2, V1, B2) MC, DV2, G	*Sar tu Rom* (MC), (R),

Table 6.8: Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

There was one variable which elicited responses from four of the seven participants, and this was standard English imperative 'shut up'. This produced five distinct variant phrase forms with one phrasal variant proximal between participants. *Kekka rokker* and proximal *kek rokkka* were elicited from Clinton and Ben respectively. Table 6.7 illustrates *kekka rokka* used by Clinton for 'stop talking'. In Matras' 2010 corpus the phrase '*kekka rokker Romanes up the gav'* ('don't speak Romanes in the town') has the same illocutionary force as the phrase used here by Clinton.

The function word *kek* also appear as part of the phrasal variant for Graham's production of *kekka chave*. C*have* is derived from the European Romani *čhavo* meaning 'boy' or 'girl'

(Manchester Corpus, 2021). This phrase translates literally as ‘not boy’ and can translate into standard English as ‘no mate’, or ’ I don’t think so mate’. Excerpt 6.3.6 illustrates the use of *kekka chave* within a narrative relating to the use of Anglo-Romani being used to covertly communicate (For discussion on ‘argot’ and ‘secret language’, see section 3.4.2. Excerpt 6.3.6).

Excerpt 6.3.6

1	Graham:	*...or you'll hear it like, I can if I go round car boots or anything like that,*
2		*I can*
3		*tell before they've even opened their mouths*
4		*and you just hear one word don't*
5		*you and you know*
6	Interviewer:	*in that situation would you probably keep quite*
7		*about the fact that you could understand?*
8	Graham:	*well I have before I mean, I've been on car boots before and buying stuff*
9		*and I've heard an old bloke say to his young lad who's eight or ten years*
10		*old who's selling stuff* ***mace*** *the* ***mush*** *of his* ***luvver*** *and I've gone* ***kekker***
11		***chave*** *and they just go like eh! that just go like you understand? oh yeah*!

Graham’s account of Romani being used for covert purpose describes a scene where he is confronted with a sales pitch at a car boot. In the example, Graham hears the use of Anglo-Romani ‘*mace the mush of his luvver’* or ‘trade the man of his money’. Graham illustrates a scenario where Romani vocabulary helps the traders communicate whilst also creating inclusivity for Graham as part of the resolution of the story. Graham arguably, creates a positive stance towards the situation of covert use as a positive characteristics.

The term *chichi* used by Jake is found across a number of sources within the archive used for ‘shut up’. The European Romani word *či(či)* is a cognate word and translates as 'nothing' (Manchester Project, 2021). The term phrase *nai puka* and *ma puka* are also single distinct elicitations from a single participant. The adverb *nai* used for negation is described in 6.1.3, whilst the variant puka is described in 6.1.1. The phrase *nai puka* translates as ‘no tell’ or ‘shut

up'. *Ma puka* is used also found within the archive sources and used in the same context as *nai*. For Ben, these variants are both used for this meaning.

The interrogative 'Are you a Roma/Traveller' produced three responses from Duncan, Jake and Ben. The phrase 'a *Rumnus gearer*?' omits all function words except the abbreviated English 'be' verb realised as 'a' as in 'Are you a *Rumnus gearer*?', meaning 'are you a Roma man ?'. The variant *gearer* is attested as being cognate to *goro* which means 'man' in European Romani (Matras, 2010). It is also attested as used in Old Welsh Romani to mean *gawdja* (see section 7.1.1 for description and variation use of *gawdja*) referring to someone who has intermarried with 'Gypsies' (Sampson, 1981). This is not the case with this use of *gearer* as it refers more generally as a term for 'man' rather than specifically for non-'Gypsy' or *gawdja*. Jake uses the phrase *pukker your nav* meaning 'tell your name' and Jake uses the phrase *sar tu Rom*. In the case of Clinton and Jake, function words are either omitted or English function words are used as with conjugation of 'be' in *a Rumnus gearer?*. The morphosyntactic rules around colloquial regional English (Matras, 2010) is shown as a characteristic of Anglo-Romani in these examples, as omission of pronouns and articles are a feature of some vernacular varieties of English. This supports the ideas that Anglo-Romani is predominantly lexical within an English matrix language but also that this is as an incomplete description. These examples support the notion that Anglo-Romani although not morpho-syntactically tied to European Romani (i.e. Anglo-Romani has omission of articles, no obligatory copulas in existential sentences and pronoun agreement for European Romani), it does operate grammatically differently from English.

Ben used the phrase *sar tu Rom.* The *sar* is a cognate of Europea Romani *san* meaning 'to be 2SG' (Manchester Project, 2021). The structure *sar tu* is also found in the Manchester Corpus (2021) although proximal in form as in *sar tutti* meaning 'how are you'. In this instance *sar* is attested to the European Romani cognate *sar* meaning 'how'. Ben's production of the verb 'to be' in *sar* is different from that within the archive in this realisation and can be attested as either an individual or within speaker group variation for this participant. *Tu* was found proximally as *tut* in the archive sources and cognates with European Romani *tut meaning 'you (obl)'*. The referent *rom* is discussed in terms of self-appellation (see section 7.1.4 for further analysis and discussion of self-referent terms). This was one of few utterances or constructions where there is no English or Anglo-Romani lexis embedded within English morphosyntax. In relation to the Ben, this characteristic is associated with his background outlined section 3.3.1 in Tables 4.1 and fig. 4.3.

No chitiaws chokas is a variant phrase elicited by Jake for 'I have no shoes'. The construction of English 'no' with Anglo-Romani lexicon is a feature of Anglo-Romani and found in other corpus (Matras, 2010). The example from this corpus shows a pro-drop tendency for Anglo-Romani as the phrase omits the functional elements 'I', and 'have'. This highlights the pattern of Anglo-Romani structured with omissions that are expected in European Romani as well as regional Englishes. See section 7.1.3 for a lexical analysis of the variants *chitiaws* and *chokas*. *Kekka* and *kek* is used by Graham on five occasions and is contextualised in the excerpt 6.3.6 and excerpt 6.3.7 below.

Excerpt 6.3.7

1 Friend: *if you heard Rye talking let's say how would she tell someone to not to say*

2 *anything*

3 Graham: *she'd probably just go* ***kekker chavi kek kek*** *yeah right shut up don't* ***kek*** *you'd*

4 *usually know just by t' eyes if she just went k****ek*** *that'd make you*

In this excerpt again Graham explained the use of *kek* and *kekker* used as a phrase to express the English 'shut up'. Here, the use of *kek* is formed with English 'don't' and shows double negation in '*don't kek'* meaning 'don't don't'. This is also a feature of Anglo-Romani reported in other Anglo-Romani corpi (Matras, 2010) as in '*Kek don't mang no kushti in her'* meaning 'don't look for no good in her' (Matras, 2010).

6. 3 Conclusion

In respect of the lexical variation analysed and discussed above there were a number of features that show a pattern within the variation elicited, along with a strong relationship with the archive sources used for the purpose of comparative analysis. In terms of individual factors that influence variation, background features indicated a patterned effect. These included regional background related to Welsh and Irish heritage in the case of Duncan and Ben, as did Duncan's proposed distinctions made between a number of variants. A number of variants aligned with archive sources for those associated with the region of the East midlands and this was seen as a general trait. This is proposed an effect related to a number of other variants throughout the chapters of analysis. A significant proportion of the elicited variants were aligned with European cognate Romani terms (Manchester, 2021) as well as with archive date sets that attest variants from largely the 19th and 20th century and corpus from 21st century sources. Individual differences accounted for the variation in productivity, and innovative

characteristics of lexical repertoire of individuals, such as with the distinction between Judy and for Jake's repertoire. Discussion of the etymological source for terms has been detailed in the case of variants where a clear source in not discernible, whilst largely either this is identified within the date or attested as a unique and as yet unattested variants. The principle that a number of terms can refer to various concepts is born out and a number of examples are given. The extent to which this relates to innovation, retention and maintenance factors as a dominant feature is outlined. Spelling and issues that relate orthography are also considered in terms of locality and arising through interactional awareness. Terms that appear more recently coined are outlined as is their ideological positioning as a cultural indexing. The position of lexis that has been borrowed into English is also examined within the interactional data and how this is positioned in terms of lexis and its ideological indexicality. The structure of phrases and there morphosyntactic characterisation is also outlined with example from interactional example with examples of phrasing where English is not the matrix language (Meyer-Scotton, 1993). Variants that show innovation produced by Jake and Ben were suggested to be characteristics of their variety. The principles behind the methodological approach were shown to have a beneficial effect in the data elicited.

In terms of interaction and performativity of identity, the use of meta-linguistic discussion was positioned as a construction of identity. Duncan used reference to ideas of word origin on a number of occasions that co-create a cultural reference positioning lexical knowledge and discussion as indexical of his own identity. Discussion of words considered by participants as 'slang' were also utilised as a means for framing identity within discussion and the alignment of terms as part of ethnic culture constructs notions of an in-group lexis and the extent to which use outside of the communities of practice aligns with an ideological sense of what constitutes Anglo-Romani or Irish Traveller Cant. The use of syntax used by Graham to show stylistic distances between those referenced in discussion and those within the interaction were also outlined as interactional indexes of identity as others also alternate use of pronouns to construct or frame positionality within interactions. Again, cultural reference to ideas are discussed. These include reference to linguistic notions such as word origin, as well as activities such as related to horses, notions that associate with outdoor and itinerant lifestyle and portrayals of morality and values with anecdotal narrative position and construct identity within discourse. The final chapter of analysis looks at variation for the semantic category 'People' looking at patterns considered thus far and again within the framework for analysis of interactionally emergent constructions of identity.

Chapter 7 Analysis: People

7.1 Semantic category four: People

The SRN 'People' consisted of four subcategorizations. These are 'Ages and Relationship', 'Body', 'Appearance' and 'Personality'. The central node around which all subcategorisation are connected were variables semantically linked to the English word 'people'. See section 3.5 for a detailed discussion of the methodology around which these SRNs were devised. The SRN elicited a total of 142 variants that participants attest as associated with Roma and Irish Traveller Pavee Cant variation that they use and spoken between in group members. This SRN elicited variants used for communicating semantic notions around referents for self and others, as well as ways of describing individuals both physically and in terms of characteristics of person and behaviour.

7.1.1 Ages and Relationship

The most productive subcategorisation for the SRN 'People' was 'Ages and Relationship'. This category contained lexis that related to words that refer to self and others, and included pronouns and words used for referring to family members and other people. Along with 'Doing things' (see section 6.1.1) this was significantly more productive than other subcategorization and considered relative to factors for language maintenance. These are arguably more likely maintained as they are used in contexts that facilitate maintenance. These domains would be expected to include for example, family and home, friendships, and employment or work (Pauwels, 2016: 90). There were a total of sixty nine variants elicited across all six participants for 'Ages and Relationship'. The productivity of this field suggests language of a minority speech community used within certain contexts of use are more frequently used, therefore more likely retained (Pauwels, 2016).

Table 7.0: SNS: People: **Ages and Relationships**

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Man	*Fin/Feen* DV4/DV4, DV5,	*Mush* MC, W3, DV2, R2, G,EA,	*Mush* MC, W3, DV2, R2, G,EA,	*Gearer/Mush* MC, DV2, DV3,D1, S1, W1, SC1, S2/ MC, W3, DV2, R2,	*Mush* MC, W3, DV2, R2, G,EA, N1,S1, W1,B2, SC1, S2	*Rom/chal/mush* MC/GY/MC,W3, DV2, R2, G,EA,

	EM, LSP	N1,S1, W1,B2, SC1, S2	N1,S1, W1,B2, SC1, S2	G,EA, N1,S1, W1,B2, SC1, S2		N1,S1, W1,B2, SC1, S2
Woman	*Bour* DV4,LSP		*Monishan* MC, S1	*Monishina/Rakkli/ Monishan* D1, F1, DV1,DV3,DV2/ MCLSC, R/ MC, S1	*Mort* MC	*Romni /Chai/ Manushi* R//DV2
Lady					*Rawni* MC, R2	
Baby	*Gorya* DV4, DV5		*Tikna* DV2, N1, FS, SC1, D1, F1, S1, B2, SC1	*Tikna* DV2, N1, FS, SC1, D1, F1, S1, B2, SC1	*Tikna* DV2, N1, FS, SC1, D1, F1, S1, B2, SC1	*Tikkna* DV2, N1, FS, SC1, D1, F1, S1, B2, SC1
Child	*Suubblin* LSP, DV4	*Chavi* IV,DV2, G,W3,GY, D1,V1,S1,W1,B2	*Chavi* IV,DV2, G,W3,GY, D1,V1,S1,W1,B2		*Chavvi* IV,DV2,G,W3, GY, D1,V1,S1,W1, B2	*Chava/Chavvie* DV2/IV,DV2,G,W3, GY, D1,V1,S1, W1,B2/
Boy				*Charva/ Chavvie* DV2/G, W1, SC1	*Chaveskro*	
Girl	*Lackeen / Youngen* LSC, DV4/		*Rackli* R, G, DV1, D1, T1, F1, S1, W1, S2	*Charva/Chavvie* B2/MC, B2	*Rakli* R, G, DV1, D1, T1, F1, S1, W1, S2	
Wife					*Monishi* DV2	

Partner	*Koura*					
Friend (mate)					*Bor (Mush,Chavali)* B2, SC1	
Non Traveller				*Gorje* DV1, DV3, LSC, SC1, DV2, R, G, S3	*Gawjer* DV1, DV3, LSC, SC1, DV2, R, G, S3	
Grandfather					*Dadeskro/Dadrus*	
Grandmother					*Daiesko dai/ Purri Dia /DV1*	
Mother				*Monishina*	*Dia* D1, T1, DV1, F1, B2, SC1	*Dye/mum* D1, T1, DV1, F1, B2, SC1
Father				*Daiol* DV2	*Dadrus* B2, AK	*Dad/Dadus* V1, F1, S1, W1, B2, SC1/DV1, SC1, DV2
Brother					*Pal/Prala/ Lakino* DV2,R, N1, R2, G, FS, SC1, S3, DV1, D1, T1, S1 W1, B2, SC1/MC/W1, S2/	
Aunt					*Bebee* R2, R, FS, SC1, S3, GY, D1 T1, F1, S1, W1, B2, L1	

I	*Meself*		*Mandi* DV2, EA, DV1, DV3, LSC, D1, B1, S1, B2, SC1, L1	*Mandi* DV2, EA, DV1, DV3, LSC, D1, B1, S1, B2, SC1, L1	*Mandi* DV2, EA, DV1, DV3, LSC, D1, B1, S1, B2, SC1, L1	*Man/ Mandi S1, W1,* B2, SC1/ DV2, EA, DV1, DV3, LSC, D1, B1, S1, B2, SC1, L1
Yours						*Tiro/tiri* G/H
Ours						*Amaro/Amari /DV3*
His						*Les/ Lesti* /EA, B2, SC1
He	*Storny (Storny korob ya – he beat you)*					
Hers						*La/lati* G,B2, S2/SC1

Table 7.1: Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.

(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

The variable that elicited variants from across all six participants was standard English 'man'. Judy produced the variant *fin* and *feen*, both documented as Irish Traveller Cant within the archive lists of vocabulary (see limitations section 3.8.2). Although Judy produced significantly fewer variants, the notion of a distinction between Anglo-Romani and Irish Traveller Pavee Cant use is further suggested through the attestations of the participants of this study. Although boundaries between Anglo-Romani and Irish Traveller definitions may be crossed or influenced by one another (Matras, 2010), the evidence here suggests in terms of lexical variation, an influence of self-ascribed identity (section 3.1.1, fig 4.3) and an awareness of indexical value (Silverstein, 2003) of these words . An example of this was Duncan and his ability to differentiate terms, and Jake and Ben's largely consistent pattern of production in terms of alignment with the archive sources. Clinton evidently produced fewer variants but also an increased ambiguity in terms of etymological heritage. This suggested some 'cross-over' but applies to limited variants produced by Clinton.

The variant used by Graham, Duncan, Clinton, Jake, and Ben were consistent with use of *mush* as a comparable variant amongst the participant that ascribe to an Anglo-Romani identity. The term relates to the cognate *murš* meaning 'man' in European Romani (Manchester Project, 2021). Clinton used the variant *gearer* as well as *mush* which corresponds with *goro*. This term is attested as deriving from older Welsh Romani and is used to refer to '*gawdja* (non-Romani) who intermarried with Gypsies' (Sampson, 1891). Clinton was not using the word *gearer* in this way and uses it to refer generically to a 'man'. As shown in the excerpt 7.0, Clinton used the term as a self-referent rather than as that found by Sampson (1891) within Old Welsh Romani.

Excerpt 7.0

1 Interviewer: *Would you say that to somebody are you a Traveller*

2 Clinton: ***Rumnes gearer** a **Rumnes***

In excerpt 7.0 Clinton used *gearer* to mean ‘person’ or ‘man’ with the self-referent *Rumnes* for ethnic identifying term. The variation produced by Ben also included the ethnic referent *Rom*, an abbreviation on *Rumnes*. Ben also used the *variant chal* along with *mush* and *Rom*. The variant *chal* is found within the archive material although this is less frequent and suggested a distinction between Ben’s background (section 3.3.1) as he self-ascribes as having a Roma heritage, and other participants who ascribe to a complex and nuanced identity (fig. 4.3).

There are three variants that provide elicitations from five of the participants. These are English variables ‘woman’, ‘baby’ and child’. The significance is the use of variants by Judy, as she produced fewer non-standard English variants but does provide significantly more lexis for ‘people’ variables. Significantly less productivity of variants is observed for Judy. However, the lexis related to the sematic category ‘People’ by contrast are still maintained within Judy’s repertoire. This indicates a significant contrast for these lexical items.

The variant *bour* has a number of spellings and within close proximity to *bour* (*beor, by-ewer, bewer, bor, bewr*). These are comparatively similar (see section 3.5 for discussion of methodology). The word means woman or wife and is attested as Irish Traveller Cant by Dawson (2002) and referenced as ‘slang’ or vernacular (Manchester Project, 2021). This production is distinct from the other Anglo-Romani participants following a pattern relatable to identity and background measures outlined in section 3.3.1, Table 4.1. 4.2 and fig 4.3. Participant’s Duncan, and Clinton, used the same variant *monishan*. They are both from the same county which may indicate a geographical influence, although not relating to individual regional identity with reference to Duncan’s self-ascribed identity (Fig 4.3). The term *manušni* means 'woman' in European Romani and several variations exists within the archive lists. Clinton and Ben also producd *monishan* and *manushi* although there is nothing to suggest these are used in distinct ways. Excerpt 7.1 below illustrates Clinton’s understanding of the use of varying lexis for the same meaning.

Excerpt 7.1

1 Clinton: ***monishna** erh **rackly** that means woman as well*

2 Interviewer: *and how would you use that differently is it interchangeable*

3 Clinton: *well if you's talking and she's got something on unusual she'd got some*

4 *unusual shoes on you'd say dick ta the* ***monishan's chokkers wi kallar***

5 *that means look at the womans shoes here*

6 Interviewer: *and then that would be different to* ***rackly***

7 Clinton: *well it is but it means the same thing it means the same thing a woman*

8 *or a female you know it means the same thing there's a lot of words what*

9 *you use what means the same thing*

In this excerpt, Clinton explained the use of *rackly* and *monishan* as interchangeable semantically. This would need further research to establish a distinction in pragmatic meaning. The use of the word indicates a blending of meaning influenced by European Romani use of *rakli* (rakkli) meaning 'non Gypsy girl' (Manchester, 2021). In this example, Clinton also overtly explained how *'a lot of words what you use what means the same thing'* showing an awareness of language change over time as a characteristic of the dialect. In this example also, Clinton used the inclusive pronoun '*you*' to refer to users of the dialect. Clinton took the position or affective stance that use of extensions of lexis is common practice for speakers of the dialect. This supported his position with the conversation as a speaker of the variety within the context of the conversation but creates a stance showing inclusivity with use of second pronoun 'you' opposed to 'we'.

Ben used the equivalent *chal* (*chai* for 'man') for woman. This did not appear within the Anglo-Romani lists and suggested a characteristic of Ben's Romani repertoire and relative to his background (section 3.3.1). Ben's use of *Romni* for women is found within the archive lists and again showed Ben's wider lexical repertoire. Jake used the variant *mort* found within the archive data and used for both Romani or specifically for non-Romani women (Manchester Corpus, 2021). There is no European cognate associated with this variant and a feature of this variant is the pejorative use shown in an extract from the Manchester Corpus (2021) below in excerpt 7.2.

Excerpt 7.2

1 Anonymous *I could* ***kur*** *that old* ***mort*** *over there*

In this example, the use of *mort* was used in a negative context. The use of gender associated terms and their contexts of use are both negative and positive within the archive sources (Manchester Corpus, 2021). An understanding of gender and identity in terms of dialect use

and specifically lexical usage would be a valuable extension to the understanding of use of dialect and how meaning and identity is negotiated through interaction.

Both ‘child’ and ‘baby’ elicit variants from five participants. Judy produced variants that correspond with words from the archive lists for the Irish Traveller Pavee Cant dialect lists and further provides evidence of a distinct pattern of use between these dialects for this participant. The principle that speakers of Irish Cant and Anglo-Romani are blending varieties together was not demonstrated by Judy. This suggested there is a clear distinction for speakers that identify as having a Gypsy, Roma or Traveller heritage. This also corresponds to Duncan’s meta-dialectal commentary who identified as having both Irish and Romani heritage. Judy used the variant *gorya* for variable ‘baby’ which corresponds with variants *gaulya*, *golya* and *goya* (Dawson, 2009; 2011). Again, these spellings represent a proximal variant (see section 3.5 of the methodology. Duncan, Clinton, Jake, and Ben all produced the variant *tikna* with the spelling *tikkna* in the case of Ben for ‘baby’. This variant is found as *tikno /tiknau/* across all lists rather than *tikna /tiknæ/* for English ‘child’. There is a tentative suggestion of phonological distinction, although as a non-codified language this difference would need further investigation. There are significantly fewer Anglo-Romani attestations for ‘baby’, however. In terms of the difference between *tickno /tiknau/*, found in the archive and *tikna /tiknæ/* it seems less likely that the ending is a spelling variant over phonological as /o/ implicates a rounded vowel based on intuitive and what are general patterns of mapping of orthography to articulation. *Tickna* is suggestive of an open final vowel, but again this would need further clarification. In terms of explaining this variation, as Ben also produced this spelling, geographical influence is not a convincing explanation for this variation. The variant *tikna* may be a more recent adaptation to the variant and is found across Anglo-Romani speakers generally over *tikno* for ‘baby’ and has not been identified thus far from previous lexical studies, partly as previous studies have not focused on variation between dialect speakers within the Anglo-Romani speaking communities. The European Romani word *tikno /tiknau/*, means 'small' in English and is phonologically comparable with the variants within the archive lists in terms of the final vowel quality and is represented by the participants of this study as varying in final vowel quality. In excerpt 7.3, Clinton explains the use of *tikna* as referring to English ‘baby’.

Excerpt 7.3

1 Clinton: *I'd say **tikna** would mean baby you'd..[hold in your] arms you know what I mean*

2 *then **charva** means six or seven year old or something like that you know what*

In this example, Clinton gave context to the different ages that the terms for children apply to. In this case, *tikna* means a child you can hold in one's arms, whilst *charva* means a child of 'six or seven'. The use of the word for 'baby' is clearly distinct from the European Romani word for 'small' here and explains the adaptation from *tikno* (European Romani) to *tikna* (Anglo-Romani for 'baby').

The variable 'child' elicited proximal variants for Graham, Duncan, Jake and Ben who produce alternates of *chavi* in terms of spelling and non-specific phonological detail. Within the archive lists there is a split between use of *tikna* and *chavi* used for 'child'. The participants within this study, however, distinguished the term tikna for 'baby' and *chavi* for child whilst the European Romani word *čhavo (/-i)* for English 'boy (/girl)' is considered a cognate (Manchester Corpus, 2021). Ben also produced *chava* alongside *chavvie* which aligns with Clinton's use of these and shows in the case of Ben a wider range of lexical variants for 'boy' and 'girl'. Ben showed to have a wider repertoire in terms of Anglo-Romani, as well as a lexical store that is significantly greater in depth and relative to his reported network (Milroy, 2002), of speakers (Table 4.3) exposure during childhood (section 3.3.1) and continued usage (see section 3.8.2 on limitations of analysis related to social network). Whilst this variant is limited to Dawson's (2002) Derbyshire list, this is not related to diatopic (by location) variation as Ben ascribed a complex regional background. This variation is considered to be due to individual background (3.3.1) related to active participation in proactively promoting Romani heritage and language together with his childhood experience.

As the participants were asked for words that mean the same as the English variable (onomasiology) the data shows *tikna* and *chavi* are not necessarily used indiscriminately. These elicitations are then compared with corpus and archive data, uncovering what these words can mean (semasiology). This analysis showed *tikna* can be used for 'child' (semasioliogical), whilst for the participants, when a distinction is needed, *tikna* for 'baby' is elicited. Previous studies have not used a simultaneous approach to lexis and specifically to variation between individuals. Similar studies have potentially not been in a position to make observations of these patterns of consistency. Judy produced a single elicitation of variant *suublin*, which corresponds with Irish Traveller Pavee Cant dialect found in the archive data set. This is a suggestive of Judy's repertoire distinguishing her dialect production from those that ascribe to Romani identity.

The pronoun ‘I’ produced six variants from across five participants. The personal pronoun is derived from the European Romani *mande* for 'me (OBL LOC)’ (Manchester Project, 2021). Ben also used *man* and this is attested across several archive sources but less frequently appearing as the abbreviated form of *man*. This variant shows as uniform across Duncan, Clinton, Jake, and Ben and is a variant that highlights another standard feature across participants both diachronically in relation to the archive material, and diatopically in relation to the locality of participant and their associated background factor for those relating to a Romani identity (Fig 4.3). Judy used a self-reflexive non-standard English form *meself*. Self-pronouns are found in the literature for Pavee Cant and again the use of these was not elicited demonstrating a narrower dialect repertoire and in terms of dialect variation, a lean towards a less clearer distinction between Pavee Cant and vernacular English.

The variable ‘girl’ produced variants from four participants from the subcategory ‘People’. The variant ‘girl’ produced more variants than ‘boy’ which could related to the use of *chavi*. The word *chavi* is described in excerpt 7.4 below and shows the distinction of use not shown through observations from the archive lists which indicates the variants *chavi* and *rackli* are used for both ‘girl’ and ‘boy’.

Excerpt 7.4

1 Duncan: *uhm child is uh **chavy** and girl is uh **rackly** okay*

From the archive lists it appeared the terms are used interchangeably. However, this does not appear to be the case for Duncan or for Clinton. Instead, it shows *chavi* is used as a general term for child as well as for ‘boy’, whilst *rackli* at least for Duncan and Ben, was used for ‘girl’ and for Duncan exclusively for ‘girl’. In terms of methodology, without the cross reference of interview, word use and comparison of archive date, this observation of variation in use would not be recorded or therefore observed. The term *rakli* is European Romani for 'girl’ although it implies non-Romani or ‘Gypsy’. Ethnicity is not suggested as referenced by use of *rackli* by the participants for this study, meaning it is potentially not maintained in Anglo-Romani for these speakers. The variation within Anglo-Romani made through this analysis shows a number of divergences from the branch(es) of European Romani. Both participants Duncan and Jake produced *rackli*, whilst Clinton explained in excerpt 7.5 below the use of *charva* and *chavvie*.

Excerpt 7.5

1 Clinton: *you just said son if it were a young son it'd be **tickna** or **charva** yeah…so if its*

2 *a girl or boy it'd be **charva** or **chavvy** you know what I mean if's a girl or*

3 *boy…its really the same word yeah*

In this excerpt, Clinton elaborated on the use of terms for 'baby', 'child', 'boy' and 'girl. Clinton explains that both *charva* and *chavvy* are used for 'girl' or 'boy' and that there is no distinction in the use of the two terms. This aligned with Ben's use of the variants for child, whilst a more detailed understanding of the use of this term is illustrated from the interview shown from this excerpt. This illustrated how lexical lists alone may not uncover such features, as the use of *charva* and *chavvie* as documented by Ben for 'child' alone does not allow for such observations. As a result, the variant *charva* is not recorded as used for 'girl' in the corpus or archive data, whereas excerpt 7.5 illustrates the terms is used for 'girl' as shown by Clinton and Ben's use.

Judy produced two variants with non-standard English *youngen* and also *lackeen* were found within the archive data associated with Irish Pavee Cant dialect. As well as suggestive of a distinct dialect from Anglo-Romani, another feature of Judy's dialect repertoire was the mixing of non-standard English and Irish Pavee Cant. A distinction between words that Judy attested as representative of her Irish Traveller heritage and non-standard English is less clear. In contrast, those that ascribed to Roma identities produce a distinction between Anglo-Romani variants and non-standard or vernacular English. This may be suggestive of a difference in how the two varieties are used by speakers but this would need further investigation. There is some discussion in relation to the use of Irish Traveller Pavee Cant (O Bail, 1994) as a register and as a creole or pidgin (Hancock, 1984) although not pertinent to the present analysis. Judy mixes both English and words found within archive sources as Pavee Cant which distinguishes her repertoire and reflects her use of Pavee Cant mixed with non-standard English. Although beyond the scope of this study, this may support a theory of Cant as a language developed through processes (Mayer, 1909) and therefore manifests as adopting or borrowing of Irish and English words along the way. In relation to discussion in section 2.2.2.2, this contrasts with the idea that the variety is a result of language mixing (Grant, 1994) but rather manifested in a perception of a variety as a register. A more progressive emphasis is to distinguish this lexical variation as a 'process' rather than the labels 'register' and 'slang' which connote a code within a language rather than a variety. This would explain the ambiguity expressed through both non-standard English and those lexical items that are found in the archive data such as *lackeen* for 'girl'.

Two variables ‘mother’ and ‘father’ elicited variants from three of the participants. In relation to participant productivity, it is the same participants that produce the variants for both ‘mother’ and ‘father’. This indicates exposure during childhood as a factor relating to Anglo-Romani dialect productivity, whilst Judy, Graham and Duncan, as reported through interview (section 3.3.1) had less exposure to Anglo-Romani. Background data shows Duncan would not consider himself to have been raised with a proficient level of dialect use, whilst Clinton and Ben report anecdotally their childhood exposure to Anglo-Romani whilst Jake did not comment on childhood experience. From the elicitation of variants for ‘mother’ and ‘father’ and the width of variants across the sample for Jake (Table 4.0), it seemed likely a greater childhood exposure to the dialect would explain the greater frequency of variants, and more specifically the semantic categorisation of variants, such as those for parents and family members. The variants for ‘mother’ and ‘father’ are all different between participants. This indicated variation on an individual level and might be expected for referents for kinships as referent terms for close family can vary between families and family members. In the case of Clinton, the variant *monishna* was used, which is also the term used for ‘woman’. In excerpt 7.6, Clinton explained the use of words for ‘father’ and ‘mother’.

Excerpt 7.6

1 Clinton: ***daiyol*** *means dad* ***monishna*** *mother*

In this excerpt Clinton expressed the use of *daiyol /dæyal/* for dad. This use was not found in the archive and corpus data. The European word for ‘father’ is *dad* (Manchester Project, 2021). The variant used by Clinton could be a derivation of the word for ‘mother’ which is *daj* meaning 'mother' in European Romani with the suffix ending *-ol* although Dawson (2002) offers an alternate explanation. The Dawson Derbyshire list (2002) also records the variant *dia* for ‘father’ and adds:

> an odd word that has been recorded in Norfolk and Scotland. The proper – but rarely heard word for father is ***dadrus***. ***Die*** being mother. ***Die*** is often used as an obscenity. ***Dia*** might be derived from Shelta ***datayr***, ‘father’, but I might be clutching at straws.
>
> (Dawson, 2002: 20)

The attestation from Clinton compliments Dawson’s recording of *dia* for ‘father’ although the phonology as transcribed for Clinton has a final velar vowel quality not suggested from the transcription found in Dawson’s record. Clinton’s use of *dia* suggested a diatopical feature of

his Anglo-Romani. This again added support to the idea that speakers of Anglo-Romani and Irish Pavee Cant maintain regional features. There may be a correlation between regional identity and regional features of dialect. Table 4.3 shows Clinton self -reports a relational score of 75% East Midlander and for Jake 40% East Midlander. This may represent a bearing on variation within the varieties of these participants although further investigation would be needed. However, within a case study analysis of lexical use, incorporating attitudinal data, and interactional analysis, a relationship between regionality and dialect is suggested. Jake and Ben use what Dawson (2002: 20) describes as 'proper' referring to a 'standard' form of Romani and recorded amongst archive sources. These are *dadrus* and *dad*/*dadus* for Jake and Ben respectively.

The variants for 'mother' vary across participants as previously alluded to. Whilst Clinton used *monishana*, this was not attested as used for 'mother' showing an individual pattern of use. Jake listed *dia* which is found across a number of sources, as is *dye*. The derivational form *die* used as an obscenity meaning female genitals was not produced, although the phonological distinction between *die*, *dye* and *dia* are not transparent from the archive lexical sets. Jake produced both *dye* and *mum*. Whilst *dad* is considered Romani derived, it is not clear *mum* is also considered an Anglo-Romani or European Romani derived term. Ben indicated that he considers this a Romani word. There are recordings of terms used in standard and non-standard English that are Anglo-Romani derived and those that may be unknown or contested by a some non-Anglo-Romani speakers. This use of *mum* is another example Romani lexis that falls within this categorisation of language that has been borrowed into use by the dominant speech communities.

One other variable produced contrasting variation between two participants and this was for English 'boy'. As discussed above, Clinton produced *charva* and *chavvie* as interchangeable variants for 'boy' or 'girl'. In terms of variation, Clinton's use of *charva* for 'boy' is found within the Derbyshire lists and alludes to the premise of variation based on region or locality. Whilst *chavvie* is recorded as used for 'boy' across various sources, *charva* for 'boy' is not. As mentioned *chava* for girl is also not found within the archive data although proximal *chavali* for 'girl' is recorded. Jake produced the variant *chaveskro* not found in the archive lists and represents an additional unattested term. This may be idiolectal (as maybe for all single participant elicitations) as well as diatopic in terms of the pattern of variation this lexical variation represents.

Fourteen of the variables for semantic subcategory field ‘people’ elicited variants from one participant. Following is each participant’s variation analysed alongside the archive and corpus data, with the most productive participant’s variation analysed first, followed by the next most productive. Jake was the most productive for the subcategory ‘People’ overall and produced eight single participant elicitations for variables for subcategory people.

Jake produced *monishi* for ‘wife’ and again is attested in the Derbyshire list (Dawson, 2002). This related to his self-reported background being from the East Midlands. Proximal variants are attested as well with all being similar to the European *manušni* for 'woman'. M*onishi* stands alone as attested from the East Midlands. For ‘friend’ Jake produced *bor* which is proximal to two archive sources *boro* and *bor* referenced respectively in the lexical Table 7.1. *Bori* is attested as meaning ‘big’ or ‘large’ in Anglo-Romani (corresponding with *baro* meaning ‘big’ in European Romani). Jake used *bori* for ‘big’, although this meaning may not necessarily correspond with the use for ‘friend’ and the distinction between *bori* for ‘big’ and *bor* for ‘friend’ would need further research to identify semantic connections between lexical variants.

Jake also used *mush* and *chavali* and alludes to these as more informal terms of referent by parenthesising ‘mate’ in his elicitation. These, as previously described, can be translated as ‘man’ and ‘boy’ from Anglo-Romani. This informal usage is also found within the archive data.

Excerpt 7.7 shows use the exonym or outside referent term as used by Duncan and illustrates its use in context.

Excerpt 7.7

1 Duncan: ***Gorje*** *which is a non-Traveller and that's as I say that's used to this day that's*

2 *used very common very very common every day you would say that word its*

3 *being used you know what I mean because we're meeting you know different*

4 *people and I would say I spoke to* ***Gorje*** *man today you know and that's a word*

5 *I would use amongst the family my children would say that very very often you*

6 *know and eh a glass of water a glass of* ***parny***

In this excerpt Clinton highlighted the collective use of the term *Gorje* and the significance it plays in the lives of family and as a community. The repetition of ‘*very common’* and again ‘*very’* shows a degree of modality in relation to the proposition ‘*every day you would say that*

word its being used'. In this excerpt, Duncan illustrated the use of the word as an in-group term and associates its use within family members and the wider community, emphasising the commonality and frequency of use through an expression of modality showing an evaluative stance.

Jake produced the exonym or outsider referent term *gawjer*, related to *gadžo (/-i)* the European Romani for 'non-Gypsy '. This term is found across a number of Anglo-Romani sources in a variety of proximal forms. None of the spellings exactly match that of Jake's own orthographic representation but are all proximal in terms of syllabic consonant and vowel patterns. This highlights the non-standardised spelling patterns of Anglo-Romani across sources (Matras, 2010: 5). Despite the prominence in terms of attestation and more general awareness of the term, the variant only appears once across the five Anglo-Romani participants. This may in part be related to a reluctance to use a referent term that refers to someone from outside the community, and in part a result of the interview conducted by someone from outside of the community (see section 3.8.2 for discussion). Whilst this interview approach holds a number of benefits in terms of methodology, this does subtract in a number of areas, such as those variants that would not usually be shared outside of the community, as well as variants that hold taboo status. These can relate to gender, ethnic statues, sexual orientation, or interpersonal values for example (see section 3.8.2 for outline of limitation of interview method). Jake contributes an additional attestation with the variant *dadeskro* and *dadrus* for 'grandfather'. Neither of these terms are found across archive sources, as only *purri dia* appears for 'grandmother' across corpus and archive sources for variables 'grandfather' and 'grandmother'. This is an additional attestation with regards to the documentation of Anglo-Romani found within the UK, and the East Midlands of England.

Jake produced the variants for 'brother' *pal* and *prala* and are found across a number of archive sources and corpus. The term *pal* is another Anglo-Romani variant found used within non-standard English as an informal term for 'friend'. This often indexes kinship along with a number of socio-functional meanings within interaction, with parallels to the use of 'dude' explored by Kiesling (2004).

The word *pal* is an example of a Romani word that has crossed over into standard English use and to some extent this moving across has drawn attention. Words that have Anglo-Romani origins have been a focus of mainstream attention with publications in local news bulletins such as Somerset live (2021) and is acknowledged within mainstream media examples such as

‘Pompey’ (Portsmouth) regional lexis (Secret Portsmouth, 2016). In addition to this Anglo-Romani has been represented in TV drama such as the crime drama *Peaky Blinders*. The historically set drama depicts characters that use Anglo-Romani lexis and potentially challenges stereotypes held towards the use of Romani and Anglo-Romani speakers in general.

Jake also used *prala*, not borrowed into mainstream English usage and perhaps related to a less frequent use across communities of Anglo-Romani. *Lakino* is also attested by Jake and this is an additional original attestation for variable ‘friend’ and not found within the archive data sources and an additional lexical variant catalogued from the current analysis.

Jake also produced *bebee* for ‘aunt’ and corresponds with European Romani *bibi* meaning ‘aunt’. There are several proximal variants that correspond in terms of syllable and consonant vowel patterns coded below in the variant Table 7.0. This represents a frequently found form of Anglo-Romani. As would be expected, amongst the lexical representation, a majority of variants will vary not by locality, but will vary by individual (Matras, 2010: 65) and by a single time and a single place (Durkin, 2012). Jake also produces *rawni* for ‘lady’, which implies status compared with the variants elicited for ‘woman’ *mort* for Jake. *Raw*ni is a term found within the archive lexical data sets and corresponds with *rani* which is European Romani for 'lady'. The term is attested as used for ‘rich lady’ within the Manchester Corpus (2021). In terms of variation, this attestation represents an additional variant used by Jake and found across Anglo-Romani speech communities within the UK, England and the East Midlands.

Ben and Jake produced twenty three variants each for variables for subcategory ‘people’. These are all pronouns and represent a greater width of lexical and morphosyntactic repertoire than the other four Anglo-Romani speakers. There is no attitudinal data for Ben but from the interview data, background factors considered in the opening section of the analysis chapter (section 3.3.1) suggest his childhood exposure and current proactive participation in Romani network activities account for this greater lexical store. Excerpt 7.8 illustrates the background in Anglo-Romani from Ben’s experience.

Excerpt 7.8

1 Ben: *My father would **Rokka** – speak in **Romanus** to me when we were out together*

2 *say doing some farm work he would tell me names of animals, birds etc but he*

3 *hardly ever spoke **Romanus** in Front of others . Even my younger Sister’s he*

From this excerpt, it became apparent that Ben acquired Anglo-Romani from childhood and from his father who Ben attests to have spoken a fully inflected variety of Romani. This suggested knowledge and use of pronouns other than *mandi* for 'me' and 'I' depends on exposure from childhood and continued exposure to Anglo-Romani in the case of the participants within this study.

Ben produced variants for English pronouns 'yours', 'ours', 'his' and 'hers'. For 'your', Ben produced *tiro* and *tiri* and are both found with the archive sources and correspond to European Romani *tiro* or *tindro* meaning 'yours'. Both of these variants were found in archive listings although these are obscure in terms of frequency of attestation and represent less frequently used Anglo-Romani although not completely abandoned, as with this attestation. The variation between the *-o* and *-I* endings in *tiro* and *tiri* did not appear to relate to morphosyntactic meaning but rather dialect variation between branches of Romani, with the *-o* and *-I* endings recorded across branches of dialect. Ben also produced similarly *amaro* and *amari* for 'ours'. *Amari* is found in the Anglo-Romani archive sources and again is rare in terms of frequency of attestation within Anglo-Romani. This variant has either fallen out of use more generally, not been recorded, perhaps resultant from methodological constraints, or has never been a salient feature of Anglo-Romani. The terms *amaro* and *amari* are found with the literature for European Romani (Matras and Tenser, 2020: 68). Ben again produced two variants for pronoun 'his' as shown in Table 7.0 and these are *les* and *lesti*. The variant *lesti* was found within the archive data (also attested as used for 'hers') whilst *les* was only found in the literature for European Romani for 'his'. For 'hers' both forms elicited from Ben were found in the archive sources for Anglo-Romani. It was speculated that Ben was using the two forms for each of these pronouns as a morphosyntactic distinction and semantic meaning is expressed through this different form. Either this, or Ben has a wider exposure to dialect branches and switches between these dialect forms. This individual difference for Ben's dialect variation was most notably highlighted through the elicitation of these pronouns.

Judy produced three single participant elicitations for variables for subcategory 'people'. These were variants *quara* for 'partner', *meself* for 'I', and *stounhe*, for 'he'. The variants *koura* for 'partner' and *storny* for 'he' were not found within the archive data or literature related to Irish Traveller Pavee Cant and represent attestations unique to this project. These attestations along with others from Judy would need further investigation to fully document the use of these

variants within context. The use of variant *meself*, further illustrates the proposition that Judy reports both features of Pavee Cant as well as vernacular forms of English as part of her dialect. This could support ideas such as O Baoill's (1994) which propose the dialect status of Pavee Cant is a process of formulation historically and the continuation of a dialect formed through process rather than a lexical or morphosyntactic system (see section 2.2.2.3).

7.1.2 Appearance

The next most productive subcategory within the SRN 'People' was 'Appearance'. This subcategorization includes items of clothing and other distinguishing features and characteristics that can be used to describe a person. There were a total of forty five variants elicited across all six participants. This was significantly higher than the following subcategories 'Personality' and 'Body' within the SRN 'People'. This indicates a higher degree of maintenance for words that refer to relationships and describing individuals. This is characteristic of the types of lexis that is preserved within an ethnic dialect that manifests within a dominant language (Pauwels, 2016). The analysis thus far, shows the case for both dialects of Irish Pavee Cant and Anglo-Romani as distinct dialects in terms of speakers reported use, with degrees of variation related to degree of exposure from childhood, reported relational attitude towards ethnic identity, reported attitude towards regional identity, time spent in locality and current use and exposure (Table 4.1, Fig 4.3 and Table 4.3).

Table 7.2: SNS: People: **Appearances**

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Well-dressed			*Array* MC	*Kelt-up kushti/Well -togged*		*Shukar ridalo Rom/Romni*
Shoes	*Brogues* (Gaelic Irish /Scotts)	*Chogs*	*Chokkers* DV2, MC, LSP, N1, S, SC1	*Chokkers* DV2, MC, LSP, N1, S, SC1	*Churiers/Chitiaws/Chokas/Chokkars* DV2*/MC/MC, SC1,/* LSP, S1, MC	*chooka's*

Scarf				*Dickler* DV2	*Dikllo* DV2, LSC	*Diklo* DV2, LSC
Ring				*Fornie* MC, DV3	*Vongushoi/Mengro/Mengri* MC//	*Forni/anggoosti/varstengarie* MC, DV3/S1, B2, S2/
Glasses	*Specs*		*Glimmers* MC, DV1	*Glimmers* MC, DV1	*Yokingras* MC	*Yok dikengari*
Clothes				*Tats/ Toggies* DV5/MC, DV1, LSC, DV4, DV5	*/Romaneskras*	
Shirt					*Gad* MG, IV, DV2, N1, R, G, LSC, S1, B2	
socks					*Olivas* DV2, MC, LSC	
Beard	*Whiskers*		*Dura*	*Fez* DV4	*Churelo* SC1	*mui bal* MC
Tall/Big					*Bis/Borri* /DV2	
Wealthy					*Barvalipen* GY, MC, N1	
Trousers					*Bulengries* DV2, EA, LSC, V1, SC1, S2	

Table 7.3: Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.

20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

Of all the variants elicited for 'appearance', one variable produced variants from all six participants. The variable 'shoes' elicited nine variants and at least one from each participant. This is accounted for as likely a term used across contexts and within communities. Within the archive sources there were several variants for this item and across the participants this is also the case. Whilst this showed a degree of individual difference between the participants, the variant *chokkers* and proximal *chokkars* (Jake's own orthographic representation) is fairly standard in terms of its form across archive, corpus and participants in terms of phonology between Duncan's 'Chokkers' */tʃakəs/*, Clinton's 'Chokkars'*/tʃakəs/*, Ben's 'chokas/chokkars'*/tʃakəs/* and Jake's 'chooka's', */tʃu:kəs/*. Ben produced *chookers /tʃu:kəs/* which has a long vowel, whilst Jake produced four variants in *churiers*, *chitiaws*, *chokas* and *chokkars*. The variant *churiers* appears in the Dawson's Derbyshire list (2002) as does *chitiaws*. Dawson (2002) notes *chitiaws* may be a cognate of Cant *chitti*. The variant *churiers* is noted as previously unrecorded in the Manchester Corpus (2021) although found amongst several families from the Derbyshire area (Dawson, 2002). Dawson speculated this may be a corruption of *choker* or derived from the Gaelic word *crudhaich*. The two variants are indicative of a diatopical feature of Jakes's Anglo-Romani. Jakes's separate production of *chokas* and *chokkas* could indicate a distinction in consonant realisation although this would need further research. The corresponding orthographic matching variants are shown in Table 7.2 in the case of Jake's production for *chokas* and *chokkars*. The elicitation of multiple variants for one variable for a single participant indicates a wider dialectal repertoire, but also an influence from a number of

sources that Jake would be exposed to in terms of Anglo-Romani variation. This shows variation between individuals will also be determined by influence from a wider exposure to Anglo-Romani speakers, influencing the production of variants between the time and place of production (diatopically, and/or diachronically). In excerpt 7.9, Clinton described the use of the variant *chokkers*.

Excerpt 7.9

1 Clinton: *well if you's talking and she's got something on unusual she'd got some unusual*

2 *shoes on you'd say* ***dick*** *ta the* ***monishan's chokkers akai*** *that means look at the*

3 *woman's shoes here*

In the excerpt, Clinton demonstrated the use of *chokkers* embedded within English morphosyntax in '*dick ta the monishan's chokkers wi kalla'r*, and translates the phrase as 'look at the woman's shoes here'. The extent to which English is utilised for functional words, using little or no morphology or syntax from Romani (Matras, 2010: 160) is a feature of Anglo-Romani more generally. However, languages that have undergone creolization may be expected to maintain morphological features (Plag, 2003). This illustrated the assertion that Anglo-Romani is largely considered a dialect consisting of lexicon embedded within English (Matras, 2010: 131). Function words '*ta*' (to) and '*wi*' (with) appear in the above example, and'the' being an English function word as well the morphosyntactic suffixe possessive 's' in variant *monishan's* which is an English morpheme. This production aligns with the proposition that Anglo-Romani is stable in terms of this characteristic as a dialect.

Graham and Ben produced variants *chogs* and *chooka's*, neither of which are found across archive sources or corpus. This again showed a degree of individual difference, related to Graham and Ben's background data. Graham's production may represent a more recent innovation amongst younger Anglo-Romani speakers, based on his age and experience (section 3.3.1), whilst Ben's longer vowel realisation in */tʃu:kəs/* may represent a more general phonological distinction in terms of accent based on his background as an émigré and less recent exposure to innovations across younger generations of speakers within the UK. This aligns with the sociolinguistic principle that older speakers may tend towards more conservative realisations compared to younger generations (Wagner, 2012, Wardaugh and Holmes, 2021). The variants for 'shoes' aligns with a number of European variants and suggested by the Manchester Corpus are *tirax, čiox* | *kirax, tirax* | *tirax* | *tijax* from Kalderaš |

Sinti | Sepečides | and Finnish respectively. This illustrates the association between the dialect branches found across Europe and with that of English Para-Romani (Matras, 2010:131) described here as Anglo-Romani.

Judy produced *brogues* for shoes which corresponds with the standard English word 'brogue' referring to a specific style of shoe with decorative perforations. In the case of Judy, the referent is for shoes generally, and found within archive sources associated with Irish and formally Gaelic origins as well as Old Norse, and Scott's English (Oxford English Dictionary, 2021, Dictionary of Scott's Language, 2021). Judy's use of dialect is distinguished as she produces variants found within vernacular English sources. Judy is less productive in her range of lexis, although Judy produces variants that show relation to Irish Pavee Cant archive sources in relation to her self-ascribed identity.

Two variables produced variants across five participants, and these were 'glasses' and 'beard'. Both of these features are distinguishing features for an individual and perhaps explains the productive prevalent elicited across five of the participants (Pauwels, 2016). In the case of Graham and Duncan, these variants are the same. The variant *glimmers* appears in the Manchester Corpus and is assumed a more common innovation feature found across the Anglo-Romani speech communities. The variant is found in in the archive across Northern regions is archaic English continued in use within speakers of the dialect. A number of non-Romani lexis is also associated with Anglo-Romani, including Cant as previously highlighted, and also old English non-standard terms that have been borrowed into Anglo-Romani as part of the dialect. This variant has been found in Northern areas, and as part of this research, within the East Midlands speakers of Anglo-Romani. In excerpt 7.1.0 Duncan described the use of *glimmers* for glasses as used by himself.

Excerpt 7.1.0

1 Duncan: *as a young person growing up in my area okay, so what we have here, glasses*

2 *glasses you said here glasses we would say* ***glimmes glimmes*** *and its a word*

3 *for looking* ***glimmer glim glim*** *you know* ***glimmes***

In this description, Duncan suggested that the use of *glimmers* is associated with speech as a younger adult as he refers to *'as a young person growing up'* in relation to the use of the word glimmers. This suggested age as a factor of use of Anglo-Romani as Duncan expresses his understanding towards age related change as an influence. Duncan uses the pronoun '*we*' in

‘*we would say **glimmers’*** positioning himself as within group in the context of the interview. There is no European equivalent for this variant. Jake and Ben used *yokingras* and *yok dikengari*, the former of which is found in the Manchester Corpus (2021) whilst the latter is a unique attestation and not found within the archive data sets or corpus. The European Romani word *jakh* for 'eye' is a cognate term and forms a part of the compound in both of these variants. In excerpt 7.1.1 below and 7.1.2, Clinton explains his use of the variant *glimmers* in relation to words used for ‘glasses’ and in relation to terms used for ‘money’.

Excerpt 7.1.1

1 Clinton: *I mean dick to the **glimmer** the **gearer’s** got on you*

Excerpt 7.1.2

1 Clinton: *nah **glimmers** it’s another word what you use for glass, **glimmer** because gold*

2 *is shiny like glass, you know what I mean, **dick** to the **glimmer** the **gearer’s** got*

3 *on, you know what I mean, that means he's got a lot of gold on*

In excerpt 7.1.1 Clinton described the use of glimmers for glasses, whilst in excerpt 7.1.2 the origin of the word relating to the adjectival meaning in standard English is explained. From this description, there is a metalinguistic awareness that the term has an origin in English. Clinton indicates the position of Archaic or old English and positions it rooted within Anglo-Romani speech. The variants of Jake and Ben are more aligned with the European forms and might be considered more conservative in relation to Romani dialect. The use of *glimmers* is found in Anglo-Romani and as such an innovation of Anglo-Romani. Judy uses the non-standard English variant *specs*. This again aligns with the proposition that Judy’s use of Irish Cant involves a process involving assimilation of non-standard forms of English as well as Pavee Cant as part of her dialect repertoire.

‘Beard’ was a productive variable and also produced variants from five of the participants. This variable, as a salient and distinguishing feature is more prevalent in terms of lexical maintenance and aligns with theory of lexis that is more widely preserved over less frequently elicited (Pauwels, 2016). The variants differ however, for each of the participants. In the case of Duncan’s production of *dura*, it appears that there is a misreading of ‘beard’ as the meaning of dura appears as ‘bread’ in the Irish Traveller Pavee Cant list of Dawson (2011), although

not an Anglo-Romani word in terms of attestations in the archive data sets. Excerpt 7.1.3 illustrates the discussion involving the production of *dure* from Duncan.

Excerpt 7.1.3

1 Duncan: *for this one here we would say **dura** I don't know how to spell it **dura dura** here*

2 *what's this now Joe let's have a look here uh **arrai** a well dressed a well dressed*

3 ***arrai***

In 7.1.3, the reference to either 'beard' or 'bread' is not clear. As Duncan described difficulty with reading and the listing of *dura* as bread within the archive data sets, it is a logical conclusion. Duncan's background (3.3.1) also corresponded with the words associated with Irish Traveller Pavee Cant. On this occasion, and being early on in the discussion, Duncan did not mention the association of this variant with Irish Traveller Cant. However, later in the interview, Duncan did comment on the difference in the dialects of Anglo-Romani and Irish Cant. As the discussion developed, the distinction between the two dialects became a central topic and assembled the idea that these two dialects are separate, as lexis is consciously produced to identify with one or both cultural identities Duncan ascribes (fig 4.3).

Below in excerpt 7.1.4 Clinton used the term *fez* and again this was not found within the Anglo-Romani lists. This was however, found in the Irish Cant word list (Dawson, 2011). Excerpt 7.1.4 shows Clinton used the term as a general reference to hair.

Excerpt 7.1.4

1 Clinton: ***fez** means hair*

The variant *fez* for 'beard' or 'hair' more generally, as with *dura* for 'beard' appears to have crossed over from Irish Cant into the speech repertoire for these participants. Further distinction between perceptions of formality and status of words with Irish Cant compared to Anglo-Romani lexis with European origins would be beneficial but is beyond the scope this research. Further questions related to formality and register between forms would be a valuable addition to the present enquiry. Jake and Ben both produced Anglo-Romani with European Romani cognates. *Churelo* and *mui bal* were both found in the archive lists, although terms for 'beard' are less frequently recorded. This was assumed due to methodological factors rather than variants being less frequent and demonstrates the advantages of the approach using the SURE method and adaptations made for this study. Churelo has a cognate in Europena Romani

čhoralo meaning 'bearded', and *mui* in *mui* for face in European Romani and *bal* in *bal* for ‘hair’. One speculative proposition is that Jake and Ben maintain a conservative Anglo-Romani, whilst Duncan and Clinton have appropriated a different register or vernacular form of Anglo-Romani. Both Jake and Ben have a differentiated register of Anglo-Romani due to factors of current exposure and use, specifically their involvement in pro-active movements involving the use of Anglo-Romani. Judy used the non-standard variant *whiskers*. This again is an example of a variant from Judy that is found used within non-standard English and corresponds with the proposition that Judy’s Irish Cant is characteristically distinct from the Anglo-Romani of Graham, Duncan, Clinton, Jake and Ben’s which has been described as largely based upon lexical insertions (Matras, 2010).

The next most productive variables were ‘well-dressed’, ‘scarf’, and ‘ring’ with variants elicited across three participants. There are synonyms for the term ‘well-dressed’, although no exact matched show within the archive lists. *Arrai* was found as rai used for ‘rich man’ in Anglo-Romani and recorded in the Manchester Corpus. This corresponds with *raj* meaning 'gentleman’ or ‘lord' in European Romani. Excerpt 7.1.5 shows Duncan’s explanation of the variant.

Excerpt 7.1.5

1 Duncan: *let's have a look here uh **arrai** a well dressed a well-dressed **arrai**... we would*

2 *say if we'd seen a big man coming down and he was well dressed and he was*

3 *well groomed he's **arrai**...*

In this excerpt Duncan explained how this variant is used to express someone’s attire. This is not recorded within the archive used specifically this way. This again emphasises the methodological approach adopted and adapted for this study and highlights the contribution of attestations recorded by this study. Lexis as an in-group practice is positioned within the interaction with the use of the pronoun ‘we’ and is relational to Duncan’s self-ascribed identity (fig. 4.3). The exchange implicates Duncan’s own identity within the interaction as being part of a larger community of speakers (see section 3.8.2 for limitations). The variant *well-togged* although not found in the archive, has an associative Irish variant *tugs* recorded as Irish Cant in the Recordings found in the Leeds Special Collection (1979), and proximal *togeri* is also recorded for Anglo-Romani (Dawson, 1959-69). The use of *kelt-up-kushti* is not found within the archive and is an unique attestation for this study. Again, these variants show a mix of Irish

Cant and *kelt* embedded within a prepositional verb phrase, indicative of possibly an informal register. This aligns with the proposition of Clinton's dialect as distinct from that of Jakes's and Ben's lexical repertoire and production. Ben produced more formal lexical variants in terms of their European Romani associative forms. Ben produced *Shukar ridalo Rom* and did not appear in the archive data sets for variable well-dressed, however, *shukar* for person did appear in the East Anglian (1929) data set and proximal *ridopen* for clothes also appears (Calso, 20th C) and can be associated with the European Romani *uriben* meaning 'clothes' in European Romani. Ben was proficient in his ability to assemble phrases without embedding Romani within English and demonstrated a wider as well as greater depth to his repertoire for his dialect as well as a more conservative or standard lexical field. In excerpt 7.1.6 below, Clinton explained the use of phrases for variable 'well-dressed'. In this instance, Clinton was using the word *gearer* not used by Ben. I would argue this variant is part of a regional Anglo-Romani dialect as it is not associated with European Romani, as is the case with *toggies*. This variant is found within Irish Cant sources.

Excerpt 7.1.6

1	Clinton:	*no you'd say dick to the **gearers toggies***
2		*he's **keld** up **kushty toggies** means clothes*
3	Interviewer:	*I'm interested in how some of these words have gone into local slang as*
4		*well I mean this isn't slanf but alot of people might say*
5		*this is Newark slang*
6	Clinton:	*it ain't Newark slang it's come from Gypsy people but some of the words*
7		*is Gypsy people's and some of it what they've added on*
8		*you know what I mean*

In excerpt 7.1.6 Clinton described not only the use of variant phrases for 'well-dressed', but also outlines his position in terms of the status and use of dialect variants within his repertoire of Anglo-Romani. He contested the idea that some words have been adopted as 'slang' within the local community and testifies the assertion that these lexical variants are 'from Gypsy people' and therefore belong in this respect to this community. The acknowledgment that *'some of it what they've added'* is also part of Anglo-Romani and that this too is part of the variety

also reinforces the idea for Clinton that Anglo-Romani holds a unique status distinguished in part to that of the European dialect branches.

‘Scarf’ also produced three variants from three separate participants. This is elicited as *diklo* for both Jake and Ben, and *dicklo* for Clinton. These three variants appear in both the Derbyshire list (2002) and the Leeds Special collection recordings (1969). These are the only recorded attestations for ‘scarf’ and not enough data to ascertain the diatopical characteristic of these forms. The observation that these forms are diachronically stable can be stated however, as they remain aligned with the realisations recorded from the 60s and early 2000s. The current methodological approach enables this observation as archived historical data gives context to the production of variants. *Dik(h)lo* is recorded as meaning '(hand)kerchief’ or ‘shawl' in European Romani. ‘Ring’ also produced variants from three participants. Duncan again produces a variant associated with Irish Cant with *fornie* recorded as *fawni* in the Manchester Corpus (2021) and attested as Shelta. It is also recorded amongst the Derbyshire Travellers although appears in Ben’s repertoire so may not be a diatopical feature of Anglo-Romani dialect within the locality, but rather a feature of Clinton’s repertoire of Anglo-Romani and relates to his dialect repertoire and background. Jake produces three separate variants. *Vongashoi* appears as proximal *vongusha* found in the Manchester Corpus and associates with European Romani *vangrusti* meaning 'ring'. The variants *mengro* and *mengri* don’t appear in the achive lists as ‘ring’ although *-engri* does appear as a suffix on *kannengri* meaning ‘earring’. *Mengro* and *mengri* are assumed alternate suffixes for *vongushoi* and both used for ‘ring’. These attestations again are not found within the archive data sets and are additions recorded through this research. Ben produces three variants, the first *forni* attested as associated with Cant. *Angoosti* /æŋgu:stə/ is found across three sources in proximal form. The orthography and phonology in terms of initial consonant are distinct and therefore represent an induvial difference in terms of variation. This may be a characteristic of Ben’s background, as in all the other attestations except for one (Sampson, 1891), the initial consonant is a bilabial approximant /bæŋgu:stə/ or labio-dental fricative /væŋgu:stə/. It is more likely as individual difference, however, than a stable feature as there is only one previous attestation from the 19th century amonsgt many other recordings. *Varstengerie* is not found in the archive data sets and is an additional original attestation recorded as part of this research project.

One variable elicited variants from two of the participants and that was ‘clothes’. Additional attestations not found across archive sources and these are *tats* and *romaneskras* used by Clinton and Ben respectively. As discussed, *toggies* is found in Irish Traveller Cant lexical data

sets. *Tats*, was also found in the Irish Cant word list 2009 (Dawson, 2009). This showed a degree of 'borrowing' or influence from Irish Cant in the case of Clinton. Whilst there is a conscious awareness of the separation of both Irish Traveller and Anglo-Romani dialects for Duncan, this may not be the case for Clinton. There are a number of variants that were shared or borrowed into the repertoire for at least some Anglo-Romani speakers from this case study. The degree to which these index distinct cultural identities is a complex and nuances picture, however. For Clinton, diatopical features as well as lexis with varied heritage form part his repertoire. This is not the same as Duncan's separation of the two dialects and outlines the uncertain lines between the two dialects depending on their context of use. This is illustrated within the meta-linguistic narrative explored within the analysis of interview data. Jake used the more conservative Anglo-Romani variant *romaneskras* and again represents a difference in terms of his lexical repertoire. The variant is not attested in the archive sources and used a suffix using inflectional Romani. This represents Ben's use of inflectional Romani and is more generally a characteristic of his repertoire. This again demonstrates a more formal Anglo-Romani dialect represented by Jake and reflects background factors. These factors are propositioned here as greater exposure to Romani dialect and a more inflectional dialect acquired from an early age and a wider exposure to this variety of Anglo-Romani. This is also a feature of Jake's Anglo-Romani, whose Anglo-Romani is characterized with inflectional features. This is also related to degree of exposure (although Clinton does not feature inflectional characteristics) and being part of a wider network of Anglo-Romani speakers, as reported in his background data (section 3.3.1).

The remaining five variants were produces by Jake and these were single elicitations for five separate variables. These were *gad* for 'shirt', *olivas* for 'socks', *barvalipen* for wealthy, and *bulengries* for 'trousers'. *Gad* is found across several archive sources and represents a variant common across sources in terms of locality, time and individual. There is variation within the source data, although *gad* is stable. *Gad* is 'shirt' in European Romani, and showed some lexis corresponds more closely with Anglo-Romani than other lexical variants in terms of syllabic structure. This contrasts with the patterns of variants produced by Jake and Ben that are more aligned with European Romani and composed of inflectional elements. The variant for 'sock' *olivas* was also produced by Jake and is also found in the archive data. The Leeds Special Collection recording showed *kallivers* is a proximal variant and shows a degree of stability for this variant between sources in relation to diachronology. There are few instances of this variant being recorded and again is an additional benefit of using the SURE methodology for

ascertaining diachronoligcal variation, as well as diatopical and individual changes. The European word for trousers is attested as *holova* and is a near cognate (Manchester Project, 2021). The adjective 'big' has an alternate attested spelling by Jake and illustrates the variation in spelling used by speakers of Anglo-Romani. In Table 7.2, the spelling *borri* is used, while *bawri* and *bori* are used in Table 6.3. This highlights the need and challenge for the development of a systematic method for comparing orthographic representations across study and methodology. *Bis* is an original attestation for variable 'tall' and represents an additional attestation for Anglo-Romani lexis. The distinction between adjectival 'tall' and 'big' was attested as semantically separate by Jake, whilst this was not previously recorded. *Bori /bɔɹi/* was found in the Derbyshire list (Dawson, 2002) and phonologically represents a diatopical feature for Jake's repertoire. For variable 'wealthy' there are three proximal variants recorded in Table 7.2, although the variant *barvalipen* /bɑ:vælipen/ is an original attestation in terms of phonological representation. The nearest form is *larvallapen* /lɑ:vælæpen/ recorded by Norwood (19th Century) which differs in only the first consonant and potentially the third vowel. This demonstrates a degree of stability in Anglo-Romani dialect, and the degree to which variants change with individual, time and place. European Romani word *barvalo* means 'rich' (Manchester Project, 2021), whilst *barvalipe(n)* means 'wealth' and 'richness'. Again, this demonstrates a formalised, inflectional characteristic of Jakes's Anglo-Romani more generally. The variant for 'trousers' *bulengries*, is found in proximal form across several sources noted in Table 7.2. No exact match however, is recorded. The nearest match is that recorded in the Derbyshire list (Dawson, 2002) in *buiengris*. The Manchester Corpus (2021) suggests cognate European word *bul* for 'buttocks' or 'bottom' as an origin. Again, the individual variation represented by Jake's production of *bulengries*, suggests a degree of diatopical influence in accounting for variation, as well as individual difference. Individual differences are characteristic of dialects with speakers separated in time and locality than for dialects with speakers closely networked (Pauwels, 2016).

7.1.3 Body

Lexis categorised as related to the body produced 35 individual elicitations between the six participants. These are shown in Table 7.4. This category represents a similar degree of productivity in terms of the total number of variants elicited compared to 'Appearance'. However, there was a noticeable gap in variants produced by Judy and Graham for this category. These may be accounted for due to background to the self-reported childhood exposure (section 3.3.1) ,self-reported identity scores (Table 4.2), relational identity measures (fig. 4.3) and self-

reported social network scores). Less frequently occurring lexis is not elicited within this subcategory, which may relate to individual differences regarding exposure during childhood and adulthood. Following is the analysis of the lexis within the subcategory 'Body'. The analysis gives lexical descriptions relative to archive sources, individual background factors, and discourse narrative features that considers social meaning of lexis in interaction within the meta-linguistic discussion around lexical variation elicited through interview.

Table 7.4: SNS: People: **Body**

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Mouth			*Mue* D1,S1,W1,L1, AK, IV, R,R2	*Mowei* DV2	*Mowi* DV2	*Mui/ mashka* D1,S1,W1,L1, AK, IV, R,R2/
Head				*Shurrer* R, MC	*Sheroe* DV1,D1, V1, B2, SC1, MG, DV2	*Shoonesa*
Nose			*Nook* DV1, D1, F1, S1, W1, B2, SC1, L1, R, IV,DV2, R2, G, S3	*Blert*	*Noker* DV2	*Knok* DV1, D1, F1, S1, W1, B2, SC1, L1, R, IV,DV2, R2, G, S3
Teeth			*Dans* G, S1, B2, SC1	*Danyes* DV2,	*Danyers/dand* DV2/DV3	*Danda* DV3
Ears				*Shunners*	*Kaun* B2	*Kan* DV3, D1, F1, S1, SC1, R, G, S3
Eyes			*Yok* N1, DV1, S1, V2, B2, SC1	*Yokkers* MC	*Yok* N1, DV1, S1, V2, B2, SC1	*Yoka*
Eyesight					*Dikomus*	

Hand			*Vasta* MG, IV, DV2, N1, G, EA, DV1, D1, N1, V1, F1, V2, W1, B2, SC1, L1	*Fams* MC	*Vast* MG, IV, DV2, N1, G, EA, DV1, D1, N1, V1, F1, V2, W1, B2, SC1, L1	*Varsti* G, S3
Breast					*Berk* DV2, R, SC1, GY, MG, S1, B2	
Hair				*Fez* MC		*Bal* MG, DV2, R
Backside					*Bulomengro*	
Legs,					*Airaz* DV2	
Arm					*Vasti* DV2	

Table 7.5: Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.

(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

Five of the variables produced elicitations from four participants. These were variables 'mouth', 'nose', 'teeth', 'eyes' and 'hand'. Participants who reported a greater exposure to Romani or Irish Traveller dialects during childhood produced variants for lexis related to 'body' or bodily related lexis.

Duncan and Ben produced approximate variants *mue /mue/* and *mui /mui:/* for 'mouth' and show variation in the final vowel. This observation is an approximation as transcriptions vary between archive sources. This is discussed within the limitations section 3.8.2. Clinton and Jake also produce comparable versions of this variant. These are *mowei* and *mowi* respectively, and produced as approximates in their orthography. The variants produced by Clinton and Jake corresponded with the Derbyshire vocabulary list, demonstrating a potential regional influence on the realisation of *mowei* or *mowi*. Ben also produces the variant *mashka* which is not found within the Anglo-Romani word lists. This suggested Ben was exposed to a wider 'European' variety of Romani and not only Anglo-Romani. Ben's variety has a pattern of inflection and morphology which represents broader characteristics of 'European' Romani. It appears to incorporate traits of Anglo-Romani as well as 'European' Romani features. In European Romani *muj* is 'face' or 'mouth' (Manchester Corpus, 2021) which relates to the variants for 'mouth' produced by Clinton and Jake.

The variants for 'nose' are consistent between Duncan, Jake and Ben and are all approximants related to the European Romani word *nakh* for 'nose'. In excerpt 7.1.7 below, Duncan described his experience of this word and relates it to his father's proficiency in Romani.

Excerpt 7.1.7

1 Duncan: *no problem no problem nose is **nook** as I said teeth is **dans** hand is **vaster** okay*

2 *some of these words if you asked my father he'd probaby tell you eh as I say*

3 *some of these words are **tigna** means small it doesn't just mean baby but it*

4 *means small*

In this excerpt, Duncan explained his father would be able to *'tell you'.* This suggested an awareness of the attrition of dialect and an understanding that his father would have used a greater knowledge of Romani words. Duncan showed his degree of confidence in his

explanation of Romani word by use of hedging, positioning an authority on the discussion of lexical variation.

Duncan, Jake and Ben produced *nook*, *noker* and *knok* respectively and appear similar. Clinton, however, produces *blert* which is a unique attestation amongst the archive data and is part of his individual idiolect, or an in-group pattern of variation. The variants for 'teeth' have a similar pattern of variation in that all variants produced by Duncan, Clinton, Jake and Ben are similar and cognates of European Romani *danda* meaning 'teeth'. Ben's variant *danda* was the same as the European variant, whilst *dans*, *danyes* and *danyers* were elicited from Duncan, Clinton and Jake and are all approximate in their orthography. Whilst there is variation within the archive data, the consistency between diatopically associated forms from Duncan, Clinton, and Jake suggest a regionality for these variants. European Romani for 'eye' *jakh* corresponded with Duncan, Clinton, Jake, and Ben's production of *yok /yæk/*, *yokkers /yakəz/*, *yok /yak/*, and *yoka /yækə/* respectively. These variants showed similarity in vowel, consonant, and syllabi. Ben showed a unique final vowel feature, again demonstrating an idiolectal variant feature. In excerpt 7.1.8 below, Duncan explained the words *yag, yock, pagger* and *gel.*

Excerpt 7.1.8

1 Duncan: *don't worry don't worry luck means the word luck is* ***yag ya I yagged*** *well* ***yock***

2 *is your eyes* ***yag*** *means* ***yag*** *this yep ehm* ***pagger*** *is fight* ***gel*** *means to go*

In excerpt 7.1.8 above, Duncan explained the words for 'luck' and used phonological association recite the variants as *yagged* leads to *yock* and again leads back to *yag* and then leading to *pagger* for fight. This is a novel strategy which may show some vocabulary is less frequently used by Duncan and characteristic of his dialect repertoire.

The variable 'hand' produced approximate variants *vasta* and *vast* from participants Duncan and Jake and a near approximate in *varsti*. The archive sources show a degree of similarity for the variants for 'hand' with *vast* meaning 'hand or 'arm' in European Romani. Clinton produces *fams* for hands and is attested as Cant. The inclusion of Cant in Clinton's repertoire is a characteristic feature and shows a degree of mixing between dialects of Romani and Cant.

Two of the variables elicited variants from three of the participants and these were variants for 'head' and 'ears'. 'Head' produced variants *shurrer /ʃʊɹə/*, *sheroe /ʃeɹəʊ/* and *shoonesa /ʃu:nesə/* from participants Clinton, Jake and Ben respectively. These variants share fewer phonological features although they share the same cognate *šero* meaning 'head' in European

Romani. Duncan's variant *shurrer* was limited to two archive sources that showed approximate features, whilst *sheroe* showed a greater correspondence, and *shoonesa* for Ben showed no corresponding features compared to Anglo-Romani lexis. This degree of variation is considered in relation to the background factors outlined in section 3.3.1 and the relational identities reported by the participants in Tables 4.1,4.2 and 4.3 and in fig. 4.3. Again, Ben showed a dialect not attested within the sources of Anglo-Romani. 'Ears' produced a reverse effect from participants Ben, Jake and Clinton as Ben produced *kan* which is found across a number of Anglo-Romani sources used for 'ear'. This contrasted with Clinton's production of *shunners* which is not found amongst Anglo-Romani sources although does correspond with European *šun* meaning 'hear' in European Romani. This is a unique attestation for Anglo-Romani with *shunningro* the closest (Manchester Corpus, 2021). This attestation showed innovation in the dialectal variety of Clinton as the verb *šun* 'to hear' has been used derivationally to form the noun 'ear' or *shunners*. Jake produces *kaun* which is also found amongst Anglo-Romani sources (Borrow, 1874).

Only one variable produced variants from two of the participants and this was 'hair'. Both variants *fez* and *bal* are found within the Anglo-Romani sources for lexical variants whilst only *bal* is a European Romani cognate as *bal* means 'hair' in European Romani. *Fez* corresponds with attestations within Iris Traveller Cant sources as *faz* is attested as body hair (Dawson, 2011) and *fayzum* or *farzim* also for hair (Dawson, 2009). This again suggests an influence or crossing between Irish Pavee Cant with the Anglo-Romani dialect repertoire for Clinton and shows this is characteristic of Clinton's variation. Ben produces *Bal* which is a variant that aligns with a number of Anglo-Romani sources and a feature of Ben's dialect repertoire.

There were five variants that were single variant productions, and all produced by Jake. This suggested higher proficiency. This could relate to his self-reported participation within the Anglo-Romani community promoting Romani language and culture alongside childhood exposure to Anglo-Romani dialect. The variants were elicited for 'eyesight', 'breast, 'backside', 'legs', and 'arm'. These variants were *dikomus*, *berk*, *bulomengro*, *airez* and *vasti* respectively. *Berk* for 'breast', *airez* for 'legs' and *vasti* for 'hand' all correspond with lexis recorded on Anglo-Romani sources. These also corresponded with European cognates *brek* 'breast', *heroj* 'leg' and *vast* for 'hand' or 'arm'. However, *bulomengro* and *dikomus* for 'backside' and 'eyesight' were original attestations. Both of these variants show Romani inflectional endings to derive new terms such as in the example of the Anglo-Romani innovation *kuripel* for fight into *kurimengro,* meaning soldier (Matras, 2010). In this instance *-mengro* is added to *bul*

which is European Romani meaning 'buttocks or ‘bottom'. This is an example of lexical creativity found in Anglo-Romani and not found in more conservative dialects of Romani (Matras, 2010) and a characteristic feature of Anglo-Romani.

7.1.4 Personality

Table 7.6 shows the lexis elicited for words associated with the semantic category of ‘Personality’. This subcategorisation related to ‘People’ produced twenty seven variants between the six participants within the study. This was the least productive category in terms of variation, suggesting items from this semantic category are less frequently used. The Table showed archive sources that align approximately with the participant elicited variants. The archive source codes are given in the second Table below. The following analysis describes the lexical features, with reference to the archive sources. As with all SRN, the analysis of the lexis within the subcategory ‘Personality’ is given showing relation to lexical descriptions relative to archive sources, individual background factors and social meaning in context.

Table 7.6: SNS: People: **Personality**

Informant Variable	Judy	Graham	Duncan	Clinton	Jake	Ben
Fool	*Idjat*			*Doilum/ ragged/ divvy* MC/DV5, DV4/MC, DV2,AK	*Dinlo/Dindler* DV2/DV2	*Din/ dinilo* MC/S1, B2, SC1, DV2
Rude					*Doshvalo*	
Generous					*Tachani ziengri*	*Kamli* V2
Kind(ness)					*Tukali-pen(-ben)*	
Upper class people				*Oldstuffer*	*Pre-engro folki*	

Merchant						*bik ta kinava Mush*
Dirty				*Mockutei* MC, DV1, D1,	*Chikalo* S1, B2, MC	*Mokado/Mokadi Marime* S1, B2, SC1, MC/DV1, DV3, D1, MC
Ragged						*Choro mush*
Cunning/Sly					*Choovantas*	
Bad (person)					*Bafedi (mush)*	

Table 7.7: Archive data sets

21st Century sources →	**(MC)** Manchester (corpus) Romani Project. 2021	**(EM)** Rieder, M. 2018	**(DV2)** Dawson, R. 2002.	**(DV5)** Dawson, R. 2009.	**(DV4)** Dawson, R. 2011.
20th Century sources →	**(DV1)** Dawson, R., Dawson Vocab 1959-69.	**(DV3)** Dawson, R., Special Collection. 20th C.	**(D1)** Dawson. R.M. 20th C.	**(AK)** Acton, T. & Kenrick, D. 1984.	**(EV)** Evans, I.H.N. 1929a.
(F1) Fox, Samuel. 1926.	**(GY)** Griffiths, J. & Yates, D. E. 1934.	**(G)** Grosvenor, Lady Arthur. 1908.	**(H)** Hamp. No Author. 1929.	**(LSP)** Leeds Special Collection. 1979.	**(L1)** Lucas, J. 20th C.
(MG) MacGowan, Alan. 1996.	**(R2)** Roberts. No author. 1912	**(R)** Russell, A. 1915.	**(FS)** Sampson, J. 1926a.	**(T1)** Taylor, T. 1915	**(V2)** Various Sources 1. 20th C.
(V1) Various Sources 2. 20th C.	**(W1)** Way. 20th C.	**(W2)** Winstedt, E. O. 1948.	**(S3)** Sampson, J. 1911	**(EA)** East Anglian: No author. 1929	
19th Century sources →	**(SC1)** Crofton, B & Smart, H. 1875.	**(IV)** Irvine's Vocabulary. No Author. 19th C.	**(S2)** Sampson, J. 1891.	**(S1)** Sanderson. 19th C.	**(W3)** Winstedt, E. O. 19th C.
(N1) Norwood. No author. 19th C.	**(B2)** Borrow, G.A. 19th C.	**(B1)** Bright, R. 19th C.			

The variable ‘fool’ elicited variants from across four of the participants. Judy produces a non-standard realisation of the English word *idiot* /iʔdʒɑt/. This could be explained as Judy’s a broader distinction of what she considers her ethnolect. This suggests a pattern in the production of Judy’s variety and variants or words that she associates with Irish Traveller Pavee Cant. Jake and Ben produced variants that are related to the European Romani cognate *dinilo* meaning 'fool'. *Dinlo* and *dindler* are both found on the Derbyshire word lists (Dawson, 2002) and are produced by Jake. Both *din* and *dinilo* are recorded on the Manchester Corpus (2021) and as well as other archive sources. Clinton produced *doilum* and is attested to be Yiddish in origin from the word *goylem* meaning ‘fool’, which is also recorded on the Manchester Corpus (2021). This is a variant only found attested on the Manchester Corpus and not found on other Anglo-Romani word lists. Similarly, *ragged* is found on only two word lists and these are associated with Irish Traveller Cant. This association with borrowings from both Yiddish and Pavee Cant suggests a history of dialect contact. This may be a feature of Clinton’s variety and therefore influencing his variation. Clinton also used the variant *divvy* which is originates with the European cognate *divjo* meaning 'wild' and a term found in vernacular English.

One of the variables produced variants from three participants and this was ‘dirty’. Again, all three participants produces distinct variants whilst all show approximate features of European Romani cognates. *Mockutie* produced by Clinton and *mokado*, *mokadi* produced by Ben are associated with European Romani *maxado* meaning 'dirty’ and ‘ceremonially unclean'. *Chikalo* produced by Jake shows a different variant although also with a European Romani cognate associative word *čhikelo* translatable as 'dirty'. Ben added the variant *marime* and stresses the graded or degree of meaning in ‘*mokado –mokadi* or worse *Marime*’. The variant *marime* is a unique attestation and not found in the Anglo-Romani sources available for this analysis. Unique attestations could represents additions to Anglo-Romani variation found within the British Isles and would need further examples to verify.

Two variables produced two variants across three participants, and these were ‘generous’, and ‘upper-class’. Ben produced *tachani ziengri* which was not found across Anglo-Romani word lists and is a unique attestation. The word originates in the European Romani word *čačo* meaning 'right’ or ‘true' and *(o)zi* meaning 'hear’ or ‘soul' and indicates Jake’s Anglo-Romani retains inflectional and a lexical creative characteristic associated with the Anglo-Romani branch of a Romani dialect. The word is recorded in the sources (East Anglin, 1929) but as *tatcho ezid*. *Kamli* produced by Ben aligns with lexis within the archive sources and associates with the European Romani word *kam-* meaning 'love’ or want'. Both of these examples show

a degree of knowledge and proficiency related background factors outlined in 3.3.1 and Table 4.2 and fig 4.3. They also illustrate how Anglo-Romani speakers use innovate-Rons to shifts in form and use. 'Upper-class' produces two variants, however, *oldstuffer* represents a vernacular English variant. Jake produced *pre-engro folki* which is a unique attestation illustrating a formation strategy common to Anglo-Romani. The suffix *-engro* is a common agentive derivational strategy and shown here used together with the root *pre-* to make the word referring to 'upper-class'.

Within the subcategory 'Personality' five variables produced single variants between the six participants. These were variants for variables 'rude', kind(ness)', 'merchant', and 'ragged'. 'Ragged' and 'merchant' were additional to the original SRN and were kept in as elicited as an 'and other' addition to the category within the SRN. Ben produced *doshvalo* for 'rude' which is not found in the Anglo-Romani archive data. The suffix *-valo* is found across a number of adjectives with European Romani and the prefix dosh or *doš* meaning 'fault' or 'sin' is used here with derivational morphology to create the adjective. Again, innovations of Anglo-Romani are a feature of Jake's repertoire. Jake also used the term *tukali-pen(-ben)* for 'kind/ness'. Again, this is a derivational process for creating words through innovations made either historically or individually. Ben's use of *-pen* and *-ben* in *Tukali-pen(-ben)* for 'kind(ness)' demonstrates this use of inflection. *Choovantas* is given as 'cunning' and again, not found within the archive data sets. This is potentially a unique attestation made possible through elicitation using the SuRE methodology (Llamas, 1999) that focuses specifically on lexical variation.

Ben produced *bik ta kinava mush* for 'merchant'. Following a similar pattern, Ben produced low frequency vocabulary which is not found within the Archive lists. This is another example of Ben's inflectional ability. It suggests his lexis or linguistics repertoire is more characteristic of inflectional Romani and is distinguished from Graham, Duncan, Clinton and Jakes's. The Phrase can be parsed into *kin* 'to buy' and *bik* 'to sell' with *ta* as coordinator and *mush* for 'person' and shows the degree of innovation and derivational morphology available within Ben's repertoire. Ben also used the variant *choro mush* for 'ragged' which aligns with European Romani *čoro* meaning or 'poor' with *mush* meaning 'person'. Again, suggested Ben's width of lexical awareness relative to his background outlined in section 4.1.1.

The characteristics analysed within chapters four, five, six and seven signal Anglo-Romani is not a 'broken language' (Coughlen, 2001: 6-51) as has been attested more generally by some

accounts outlined in Matras (2010: 13).To typify the dialects or to classify the dialects of either Irish Traveller Pavee Cant or that of Anglo-Romani is to overrule and ignore those speakers for whom these classifications do not apply. The variation analysed from the participants as self-ascribed speakers of varieties of Anglo-Romani, and Irish Pavee Cant are shown here to be varied and unique to each individual whilst comparable to the lists and archival sources that showed language elicited from other self-ascribed speakers of groups and communities from an extensive range of localities and networks.

7.2 Conclusion

In summary, there were a number of observable factors that influenced variations presented by the participants. Comparisons with archival data sets have also allowed for contextualisation of those variations within a historical and current comparable background. In terms of background, the data have exhibited distinctions between lexis that identified ethnicity between the speech communities of Irish Traveller Pavee Cant and those who identified as Anglo-Romani. The line between those two varieties was not opaque. A clear distinction was overtly made by speakers. An individual's background, including exposure to language varieties, such as, Roma from childhood was an influencing factor for variation. Jake's use of variable terms for parenting, recent use, and focus on the group or network of speakers in his present circumstances, also, proved an influencing factor, as with the other participants when attributing the types of innovation, productivity and, specifically, the variations in lexis, which were used by the speakers within this research. Attitude, as measured by the Identity Index Score, was also discussed as a factor of influence . Whilst empirical research might establish that as a statistical phenomenon, from an epistemological stance and theoretical position in this research, attitude could be described as an interactionally positioned phenomenon that was measurable through observations of interaction, rather than numerically alone. In that sense, the ratings collected from identity measurements could be thus viewed as only a part of the interaction between the researcher and participants in the analysis, although they proved to be a valuable instrument. Regional or diatopical influences were, also, found to be significant, when considering variations that existed within the corpus of lexis and interactional data of an individual's repertoire in relation to the wider diasporic speaker populations. That can be reinforced by examples outlined, that demonstrated associations within the historical archival data, and through triangulation achieved with comparisons within and between archival data and participant data from outside the regional locality of the East Midlands. The use of methodology was also highlighted as an integral means for qualifying those observations. The

result of influences upon Anglo-Romani from sources of branches of European Romani was also evident, and supported by comparing corresponding variants from across archival data sources. The semantic category 'People' was also found to be an influencing factor as across the distinct categories, individuals were found to follow similar patterns of productivity, showing that to be an influence. Age was also a consideration and evident as an influencing factor for variations between the participants.

The analysis, in this section, highlighted the significance of variants in terms of comparison with archival data, in addition to, as yet, unattested variants, and how those were used by speakers within this research. The analysis examined variant features of speakers, and their associations with archival data sets. Terminology for self-referent, including references to those outside the community were outlined. Terms that demonstrated multiple definitions were also outlined, thus establishing pragmatic nuances, which suggested further research could interpret those detailed distinctions of usage as terms that were exploited for contextual purposes. That context of usage was shown to vary between participants, as with the use of *chavi* for boy and *rackli,* exclusively, for "girl". Spelling patterns were discussed, as that was a distinct consideration for the methodology and was a factor within the discussion of lexical variation within the interactions themselves. Variants that had not been recorded were examined, and could be thought valuable in terms of adding to the corpus of Anglo-Romani and Irish Traveller Cant Corpus. Those were not only of significance to the corpus, but also to distinguish the methods used here as productive and effective during the interviews as a means of elicitation. The use of lexis as an argot was also described and discussed as part of the etymology of those languages with cultural associations to Anglo-Romani and Irish Pavee Cant. In addition, the borrowing and integration of words into mainstream English and regional dialects were, also, appraised as a phenomenon associated with the language of GRT communities. Further, the notion that Cant was partly viewed as a process and how this manifested linguistically for its speakers was also assessed. The use of words, such as, "*pal"*, which have been borrowed from Anglo-Romani into mainstream English was again examined in relation to usage in indexical terms, including how they were used within discoursed that centred around meta-linguistic talk with refence as to how they were used outside of GRT social networks.

In summary, with reference to the interactional and emergent production relating to identity there were a number of examples presented within this analysis of variation within the data presented in the semantic category of "people". The use of pronouns and lexis as a means of placing identity within interaction was highlighted. The position of dialect as an integral part

of community life, and how this was represented within interactions was highlighted to show how in-group status was foregrounded within discussions. The use of language, as a discrete method of communicating was discussed in previous chapters, which illustrated how individuals' produced this within interactions as a function of dialect for communication, and, therefore, as part of the character of the in-group lexis used within socially networked groups. There existed, moreover, a notion, within the data that certain lexical items were innovations, which were not always rooted in a historical connection within European Romani etymology, but had been influenced by other linguistics sources. However, that was also a part of the identity of the Anglo-Romani, and formed a part of the unique linguistic identity of those speakers within the community. Interestingly, speakers used a number of devices to show association, but, also, including stylistic distance from ethnic group depending on the context of the interaction, such as, the dynamics between the interlocutors and topic choice for conversation. More specifically, there was an awareness that the languages of those that ascribe to a GRT identity, and, more precisely, those practices associated with the referents of Anglo-Romani and Irish Traveller Cant, discussed within those chapters, including how the languages could be viewed by future generations, were dependent on the importance of their relevance . The following section offers a summary and conclusions of the research and its findings.

Chapter 8 Conclusion

8.0 Introduction

In the conclusion of this study, a synopsis has provided a summary of the background and aims, in addition to a review of the methodology used in the thesis. Following that, the conclusion outlined the qualitative findings, with a summation of the impact in respect of the qualitative data had upon the main findings. This is followed by a discussion of the contribution that this research project has made to the field of Romani, and the study of Shelta, and variation studies of usage in linguistics more widely. The conclusion has finished with remarks concerning implications of the findings in this study and possible directions regarding future research into language of the GRT communities and linguistic variation, and, specifically, lexical variation studies of speakers of minority languages.

8.1 Synthesis of the Study

8.1.1 Background and Aims

This study sought to identify and characterise non-standard lexical variation within the dialect repertoires of individuals identifying within social categorisations such as Gypsy, Roma or Traveller within the East Midlands. It sought to identify contextual factors, both social and linguistic, that influenced variations for individuals, who self-identified as Roma or Irish Traveller. It, also, proposed a method to ascertain how GRT identities emerged as relational and as a sociocultural phenomena that circulated in local contexts of discourse (Bucholtz and Hall, 2005), including how individuals exhibited identity through stylistic choices in interactions (Bucholtz, 2012). GRT communities are an integral part of UK culture and society, and can be perceived as retaining their own culturally distinct lifestyle, beliefs (Hancock, 1992), including the specific focus of this study, language and dialect (Hancock, 1984; Matras, 2010a).

The extent to which those traditional varieties have crossed, mixed, or influenced local dialects or maintained linguistic individuality has been little understood, and substantially overlooked (Beale, 2010). This study, was an incentive to gain an understanding of the largely neglected regional dialects within the East Midland districts (Braber, 2015), with an aim to identify the linguistic features of Gypsy, Roman and Traveller dialects in the speech of individuals within this region, relative to those from outside the region, and the extent to which those features had been established. It also sought to comment on any influences that those features had on local dialects .

Consequently, areas for exploration arose over linguistic depictions from communities in their accounts. First, how those individuals defined themselves within the British Isles and East

Midlands regions, and, specifically, how they affiliated or sympathised with popular, historical, and academic accounts of Romani, Gypsy, and Irish, Scottish, English and Welsh Traveller narratives . Secondly, the accounts of individuals who identified within those communities was central to an understanding of linguistic variation. That understanding searched for, in terms of how variants were used, in addition as to what variations existed between individuals, including how individuals identified themselves , and to what extent background and attitudinal factors accounted for the characteristics of variation, which were central to an individual's identity, including how they placed that variation as part of a wider social practice.

8.1.2 Review of the Methodological Approach

An ethical reflexive approach to data collection was adopted for the purposes of this research project. This was adapted from dialectal research methods, which have been carried out in the UK, with specific focus on the fieldwork undertaken on lexical variation within the regional communities of the UK (Llamas, 1999). The approach to data collection was founded on the principles of methodology, in the tradition of sociolinguistic research, where linguistic resources gained social meaning through interaction (Bucholtz and Hall, 2005). The details of the research instruments were discussed in the methodology section (Chapter 3). A mixed methods design was chosen (Creswell and Plano Clark, 2018) to enable a measure of structured qualitative analysis relating to conversations between the researcher, and the other participants, with quantitative attitudinal ratings being reported by the participants in the study.

Central to the investigation was the focus on lexical variation that was self-reported by the participants taking part in the research project, as were the processes through which identity could be framed within interactional dialogue. Self-reported lexical variation was contextualised in relation to constructs of ethnicity and self-identity measures. That was furthered through the use of a data collection method and research design by employing Sense Relation Networks (SRN) as the primary data collection instrument. The use of the SRNs allowed for the comparability of lexical variation regionally, between individuals, and historically by using archival data. That method was developed by Llamas (1999;Kerswill, Llamas and Upton, 1999).

The selection of semantic fields with standard concept terms in the four SRNs (Appendix 1) was the result of an initial trial (see section 3.5.1), originally based from a selection of words from literary sources of Anglo-Romani origins (Anglo-Romani Project, 2021). The notion words remained static, and the lexical items edited. Standard notion words with no responses were deselected where no variant was found. Each subdivision provided space for variants of

words that were not listed in the SRN and therefore allow for insertion of words that might register variations outside those items on the SRN. The SRNs were designed, therefore, to help indicate geographical and individual variation of those, who identified as Roma or Irish Traveller in the East Midlands and to develop a conception of how this ethnic and culturally associated variation could be distinguished and characterised. To this end, one participant was not from the same region, although did not provide a means for comparison on any large scale, prompted discussion relating to variations and regionality .

Unconscious responses were promoted through associations of notion words featuring in each SRN, which increased time efficiency, with an avoidance of formal interview style questioning (Llamas, 1999) that had a significant benefit in increasing the authenticity of responses. As this design corresponded with previous dialect survey subject matters (Llamas,1999), from the groupings of questions in the Survey of English Dialects (Orton and Dieth, 1971), comparability and reliability of the results for the research findings in this study were facilitated as a means of comparative analysis of the results.

The design of the semi-structured interviews was made to gather information to explore self-perceived notions of identity and language use. The method used permitted engagement with individuals within their community setting (Cohen, 1985). The discussions focused on language, and ideas about community and attitudes within the region, in addition to, nationally for Roma and Irish Traveller groups, thereby, assisting in exploration of notions of identity.

The use of an Identity Score Index was used to acquire a quantifiable means of measuring an individual's attitude towards language and identity. The index was developed by Llamas (1999; adapted from Underwood, 1988) and further modified for the purposes of this study, with specific regard to the original devised by Underwood (1988) based on Le Page and Tabouret-Keller's (1985) theory on acts of identity. Multiple choice questions were measured as scores 1, 2 or 3 with 3 representing a hypothetical scenario suggesting the strongest feelings of allegiance to community, with a separate scale created for region, and was easily noted by the researcher. The Visual Analogue Scales allowed for measures of attitude and relations of, for example, notions of self-identity. Together, the SRNs, interviews, Identity Score Index and Analogue scales were integral to the compilation of data and subsequent analysis of lexical variation identified through this research.

8.2 Main Findings

From the outset this study looked to address three questions and took two separate approaches that were interdependently linked. The first question sought to uncover the non-standard lexical variations that existed amongst those who self-identified as Roma or Irish Traveller. In doing so it was intended to identify, with reference to variables related to background, the factors that influenced, and explain variations that were uncovered from interviews and quantitative measures. As part of the data collection, the interactions themselves were also analysed in order to reflect an account of identity as the social-positioning of self and others. These relational views being based on the premise that identity was a sociocultural phenomenon produced through interactive discourse (Bucholtz and Hall, 2005). Thus, the findings would be compartmentalised and summarised with reference to those two separate considerations.

8.2.1 Lexical Variation

The lexical variation research question and associated sub-questions posed in this study were:

RQ1: What constitutes the non-standard lexical variation, with reference to semantic meaning and phonological features, for those who identify within a Gypsy, Roma and Traveller identity, within the East Midlands?

The analysis showed that, whilst there was a similarity of variants between the speakers, it exposed a complex pattern of features that differentiated in the lexical and dialectal repertoire of the six participants, who took part in this research. There was a pattern of distinction between speakers identifying as Irish Traveller and those who identified as Anglo-Romano or English 'Gypsy'. That was attested, also, from one of the participants, who identified from a mixed heritage of both Irish and Anglo-Romani background. There were numerous lexical distinctions outlined within the data that highlighted variations between the participants within the case study and corresponding corpus, such as, the use of *screeve* used for 'car' and crush for 'go' (section 4.2.2). Whilst this was a distinguishing factor, lexical variants listed as Irish Cant were recorded by those identifying from a Roma background, demonstrating a historical complexity, in terms of lexis, retained and borrowed between the speakers of those identifying from Roma and Irish Traveller backgrounds. However, e many words could be identified with a clear etymological distinction (in '*gavver*' and '*shades*' section 5.1.1) there were examples where the etymological root of the term appeared ambiguous in terms of its association with Cant or Romani (e.g. *bahlali,* section 6.1.3)

There was a general pattern for the speaker identifying as Irish Traveller to mix both vernacular English with lexis associated within the data sets of Irish Pavee Cant. Those who identified as

Anglo-Romani did not present the same pattern of production. There were some instances of historical or archaic terms retained in Anglo-Romani but changes in meaning or relexicalized and integral to the lexis of the Anglo-Romani (see discussion on variant use for 'this' section 4.2.2). This could be, perhaps, expected of a legacy language retained within what might be more broadly categorised for the purpose of this principle as a "speech community" in the Labovian (1972) sense of the term, notwithstanding the resulting and relevant critique of that concept. A pattern of variation was discussed in terms of the speaker self-identifying as having an Irish Traveller's background. Judy's variety mixed both English and Cant words as part of her variety. That process of mixing with vernacular English and Cant words was discussed as a variety that was not independent of those elements(discussed in section 5.1.1). In contrast, vernacular forms of English were found to be distinctive by those speakers of Anglo-Romani, and more so, in terms of those lexical items, which were loan terms in English .

There were a number of terms found to cover more than one semantic concept, such as, the term *hoben* used for food or for vegetables, for instance, (section 8.2.1), whilst specific terms for subordinate items were produced less, and, therefore, not retained within the speech of the participants, such as, demonstratives (section 4.2.2) or variants for food items (4.2.1). This was consistent across other corpora (Manchester Corpus, 2021), and would be expected as a minority language. However, a number of terms also revealed a pattern of lexical broadening or extension as they referred to the same concept in the case of some words for certain participants, thereby, demonstrating a blend of those meanings associated with those items (see discussion of '*boore*' excerpt 6.9).

For inflectional characteristics, the findings were consistent with a number of other studies that showed that Anglo-Romani was largely a language consisting of lexis embedded within English syntax, integrated into English morphology. That was not always the case and a number of inflectionally productive morphemes were utilised by the participants in this study. This was evidently the case in a limited number of examples for Jake and Ben, whilst Clinton also provided evidence of this, although those were mostly fixed phrases, and, therefore, in the case of Clinton, arguably non-productive. Examples were found across a number of categories such as with 'People' (7.1) as with the distinction between *drabeskro* and *drabmenagri for doctor* (section 5.1) and for demonstratives of place: *Aka* and *akla* referencing a positional difference (section 4.2.2).

A number of function words were seen to be retained within Anglo-Romani, which varied considerably between participants. In the case of Jake and Ben there were retained concepts of prepositional variants, whilst for others this had been reduced, showing assimilation of terms for some concepts such as *kuvver* for 'that'(Excerpt 4.1.1). Wider variation was evident for some deixis terms, whilst others, such as, for 'thing' were uniform across participants (section 4.2.2). This was born out across a number of terms and supported the notion that some concepts were functionally richer for some individuals than others. That was certainly the case for those terms that were indexical to identity, such as, the term *gorje* for 'non-Romani' (excerpt 4.3) and other terms like *grai* for horse (section 4.2.3).

Also apparent was a significant number of unique and potentialy unattested variants that have been recorded as a result of this research project. These were notably characteristic of the variants found within the archival data, representing potentially unrecorded terms. Those would increase the account of variation associated with the varieties linked to speakers identifying with Roma and Irish Traveller backgrounds. Some of those attestations appeared to show interdialectal influences, and a form of innovation, supported with examples in the analysis, such as, the uses of *gammy* and *glimmers* amongst other used by Clinton (excerpt 6.3.4).

RQ2: How do individual background factors of region, age, gender, attitude and identity influence non-standard lexical variation?

In terms of background there were a number of features that showed themselves as influencing lexical variation. For Jake and Ben there was supporting evidence to show the centrality to a network of speakers of dialect possessed an overall effect (see limitations section 3.8.2). This was the case for Jake and Ben as they both indicated they were part of a larger network of speakers. This network was considered to be widely connected, and, in addition, demonstrated an active pursuit of retention and maintenance of the dialect and lexis associated with their social network (see section 3.6). This was, in part, in contrast to those speakers whose background is declared as being more localised, and who, as speakers, were currently less exposed to other speakers of other varieties or, at least, less frequently so. The influence of branches of Romani was also considered as Ben maintained an influence corresponding with Welsh Romani with certain lexical variants compared with other participants within this research.

Lexical variation was not only different in degree of productivity but also in the kinds of innovations and inflection that existed between those speakers proactively engaged with the

language. The influence of childhood exposure, and the degree to which an individual was currently centric to circles of speakers was found to be a distinguishing influence on the amount of lexis, but also in the type of lexical variants produced. In the case of Clinton, for example, productivity was comparable, but inflectional classes were distinct from those of Jake and Ben, as Clinton's variants maintained few inflectional endings, and word formations, such as, those produced by Ben and Jake (e.g. *loolo mol* for 'red wine' or *tato pani* for 'spirits' (section 4.2.1). This showed an effect of diatopical influence was still retained, whereas Ben's would be less influenced, which was found to be the case. In terms of age, whilst not observed as a significant influential feature, there were certain terms that appeared to show distinctions based on age as a factor (section 5.1.1 for discussion on generational differences)

Variants that were retained within the repertoires of the speakers in this research showed a pattern that corresponded with semantic categories that associated with locality in the case of names of places (section 5.1.2) as well as relative items, such as, certain foods or terms related to certain activities or animals (see excerpt 4.7 for discussion on *pulmingerer*). As mentioned in relation to the broader question RQ1, the identification as Irish or/and Anglo-Romani bore an influence on the lexis that was produced by each of the participants. This was the case between participants, and, in the case of one participant, identifying as both, Duncan's variety was distinguishable, as such, whilst retaining both identities and lexical distinctions. A significant number of terms were also shown to have diatopical associations, such as, the production of *livna* for 'beer' amongst others (see section 4.2.1). That finding was not necessarily expected but showed nonetheless that a small number of case study participants by using a fine-grained method of detection could be utilised to find examples of regionally influenced variations.

Variant retention was supported to be an influencing factor in terms of variations , as examples from the interactional dialogues demonstrated speakers actively attempting to recall certain words that were less frequently used, and not being able to recall a certain number of variants (excerpt 6.2.8). That, whilst measures of attitude and relational measures were associated with productivity relative to self-reported background factors, such as, attestations of childhood exposure, including active participation in language practice.

Cultural saliency was also an influencing effect. A significant degree of terms elicited across participants were associated with ideas and concepts that are regarded as culturally central. However, some semantic areas were more opaque in terms of their abstraction and cultural

relevance were found to be less productive overall. The four central semantic categories of the Sense Relation Networks showed a pattern of production that varied between categories demonstrating semantic category as an influential factor for lexical variation.

There was a general pattern of correspondence with the archival data sets, in as far as a majority of the elicited lexical items for this research agreed with items from historical sources. This was borne out of the method for lexical comparison between data sources, shown to be an effective method for the purposes of this kind of historical comparative research. That revealed a significant degree of retention of lexis and dialect in terms of the varieties that were studied as part of this research. There was, however, a consistent pattern of phonological deviation from historical sources in the form of vowel or/and consonant differentiation and this pattern was consistent for lexis across this corpus. Certain examples showed themselves as consistent across locality, and this was used to support the claim of a diatopical influence or effect such as with *mushkerer* for 'police' (see discussion for excerpt 5.1) and *Mowei* for 'mouth' (section 7.1.3) .

A significant number of variants were found to correspond with archival sources and a pattern of innovation and change and development was supported by evidence. Terms, such as, *mingerer* for 'police' and historical counterpart in Welsh *prastamangra* (Sampson, 1923) was an example of this retention whilst progression was evident. Archival sources showed relation to a majority of earlier sources related to 19th and 20th century Anglo-Romani sources, with Yiddish (section 7.1.4) in addition to archaic references to vernacular English as sources of origin.

In terms of its typology, historically it could be compared to languages that were defined by their 'secret' or function as a discrete form of communication. This was not born out as a consistent characteristic of the language in terms of use, although discussed as a topic within the language as a function of identity (see discussion on excerpt 6.2). The association with 'slang' and the comparison with informal or other non-standard varieties was also discussed in terms of identity and how it was perceived and portrayed by its speakers (see discussion on excerpt 4.1.4).

8.2.2 Identity and Lexical variation

The research question related to lexical variation and identity as a social practice posed by this study was : -

RQ3: What is the social meaning of non-standard lexical variation for those that self-ascribe to a Gypsy, Roma or Traveller identity?

Identity, as an interactional and intersubjective construct, was evident through discourse and discussed throughout the data presented within the chapters of analysis. The use of an ethical reflexive approach brought into view social meanings that those individuals invested in their practices. These meanings were found to emerge in local contexts, and were placed within wider social categorisation, but, more specifically, within a local and more flexible and nuanced positions of identity (Bucholtz, 1999; Bucholtz and Hall, 2005; Jones, 2011; 2019; 2020), such as, what it meant to display knowledge of words that were culturally relevant or to offer reference and saliency to certain topics. Identity emerged in relation to explicit and implicit indexical processes. Those included overt mention of identity category labels. Identity, as constructed through interaction, was also shown as a process through which individuals positioned their own usage with respect to others. Lexis and linguistic structures that were ideologically associated with macro-social categories were analysed, as were the roles and personas within those interactional encounters. That promoted a description of identity as positioned within interactional meanings (Bucholtz and Hall, 2005). As outlined within the methodology, identity is found intersubjectively formed through relations of similarity and difference, genuine or artificial, and premise of authority, including de-legitimacy and understanding, as mentioned in the introduction, identity and its representations were constituted by the context of the interaction, were inherently relational and, therefore, partial.

Throughout the analysis were examples of situationally constituted identity practices, such as, the use of pronouns to demonstrate relationality between speakers and their roles within the interaction. Speaker "stance"(Du Bois, 2007: 163) was discussed in terms of "evaluation", of speakers' "positionality", and attribution of position to others . Other examples displayed the use of lexical and other linguistic structures in association with the ideological persona of individuals in connection with a wider social network of speakers. For example, how the use of *Gorje* meaning "non-Roma" was discussed within interactional discourse to position speakers within a coherent network of speakers and culture (see excerpt 4.3). Examples also included the use of saliant topics that were overtly indexical to cultural relevant activities (excerpt 4.1.5), in addition to, and presupposed notions of identity through discussion of lexical etymology (excerpt 5.0), including ideas and values associated through narrative and anecdote (excerpt 6.3.5), with the positioning of lexis as a means for indexing identity (see discussion of excerpt 4.1). Therefore, where speakers contested the use of certain terms "comes of us"

for instance (see excerpt 5.2), was an example of identity positioning in relation to the use of lexis.

8.3 Contributions to the field of Sociolinguistics and Language Variation

This section outlines the contributions and associations this study has made within the field of Romani studies, the study of Traveller Englishes, and, more widely, the sociolinguistic field of language variation. This section discusses those contributions from both a theoretical and methodological standpoint. Practical contributions were addressed together with future research directions in section 8.5.

8.3.1 Theoretical Contributions

The study has contributed to the understanding of linguistic variation and specifically lexical variation as spoken by speakers of a minority variety within the UK. The varieties studied within this research were spoken by individuals as part of wider social networks. It has contributed in two ways: a) it deepened our understanding of the lexical variation spoken by those who identified with Roma or Irish Traveller backgrounds, an area of variation that has been often overlooked within the region of the East Midlands (Braber, 2018), and in terms of lexical variation, a variety underexplored (Durkin, 2012; Beale, 2015) and, b) its contribution to additional knowledge of the understanding of the co-constructed nature of spoken identities (Bucholtz and Hall, 2005) of individuals of Roma and Irish Traveller backgrounds.

The main aim of the study was to explore the lexical variation of individuals that identified as having a Gypsy, Roma or Traveller background and to account for factors that influenced this variation. The literature on sociolinguistic variation did not account for regional variation within the East Midlands generally (Braber, 2015) and, specifically, of those who self-identified with Roma or Irish Traveller backgrounds. Nor did the literature account for minority varieties within the East Midlands more widely. The lexical variation examined within chapters 4 through to 7 offered a detailed and fine-grained account of that variation. In addition, that variation was contextualised within the archival data sets selected for the purpose of this research. That contextualisation has enabled a comparison of lexical variation both over time (diachronic) and across locality (diatopic). That has prompted a fine-grained analysis that accounted for lexical variation and factors that influenced that variation.

In addition to the analysis of lexis and its contextualisation across time and locale, the individuals, including factors that reflected identity as a social practice were integrated into the analysis of this research. Studies of lexical variation within the literature of sociolinguistics and identity as social practice have yet to be conducted in this way. This approach accounted

for the interactional discourse and discussion of lexical variation itself as well as in lexical variation, as a means for constructing the identities of the individuals, who took part in this research. Identity as a social practice (Bucholtz and Hall, 2005) understood from the view as speakers co-constructing identity as an approach to the study of Romani and Traveller culture has provided original contributions to these fields of research. In doing to, it has provided a detailed understanding of what GRT culture meant to individuals, who identified within those broad social categorisations, and the linguistic heritage that existed within the speakers as individuals, and how those individuals were positioned, and placed themselves as members of a minority group of speakers in relation to the East Midlands region of England.

8.3.2 Methodological Contributions

The first contribution in terms of methodology was to adapt the research tools from the SuRE research method (Llams, 1999). This research package was originally designed for use amongst a regional population that identified within the local and regionally dominant speech community (see limitations section 3.8.2). The SuRE research tools have been used by other projects (e.g., Sandow, 2020; McCooey-Heap, 2020), whilst not for those speakers ascribing to GRT backgrounds. An understanding of sensitivities that could arise, such as, mistrust of authority were carefully considered. However, any assumptions were not presumed, with the methodology being adapted correspondingly: factors that might have erected a barrier to conducting a variation study of this type (Bonevski et al., 2014), although it is acknowledged nothing should be based on supposition alone. In addition, the use of qualitative measures were also modified both in terms of content for items across all data collection tools, as was the theoretical approach to the interpretation of these (section 3.8). The theoretical approach taken within the research was based on the principle of language and variation as a socially constructed identity practice that was interactionally manifest (Bucholtz and Hall, 2005).

Further to the adaptations to the research tools, the use of data within the archive , the selection, and the practice of comparison between data sets, which were not designed for the purposes for which this research assigned, was also an original contribution to the field of Roma and Traveller studies, as well as to the wider field of sociolinguistic variation (section 3.5.3). The methodological approach defined here, utilised a best fit approach, whereby lexis was similarly based on intuitive judgements allowed for a consistent comparison of lexis across data sets that inherently used inconsistent transcription methods, or simply no system had been defined, or different transcription methods were used between different sources of lexical data. The approach undertaken by this research was pragmatic, nonetheless, it proved to be a satisfactory

method for the purpose of this type of research. In addition, previous studies had not used a simultaneous approach to lexis and specifically to account for variation between individuals (see discussion of *tikna* excerpt 7.3).

The approach for purposes of analysis, using both lexical comparison from an original collation of data sets, and transcription of interactional accounts of lexical variation (outlined in section 3.4), as discussed within the frame of identity and culture (outlined in section 3.8), provided a unique and exclusive account of the lexical variation used by those who identified with Roma or Irish Traveller backgrounds, including how the identity of the individuals was constructed within discourse. This account of lexical variation, integral to their identity, was unique with respect to the findings in this research, and could provide a template for other research that might investigate similar individuals and in-group practices.

8.4 Practical Implications and Directions for Future Research

Sociolinguistic studies of linguistic variation and, more specifically, lexical variation within the East Midlands has been underexplored and largely ignored with some significant exceptions (Braber, 2015; 2018). Durkin (2012: 3) stated: "variation in the lexicon: the Cinderella of sociolinguistics". The current study has attempted to uncover the richness of linguistic variation that has been evidently found within the speakers of those that identified as having a Gypsy, Roma or Traveller backgrounds. The use of the SuRE (Llamas, 1999) data collection methods have been integral to the findings of this research. The approach also utilised the principles outlined by Bucholtz and Hall (2005) that exposed the way in which identity was a sociocultural product of interaction, rather than an existent state within an individual. The negotiation of identities in this way was orientated within language choice and within conversations coalesced around metalinguistic discussions concerning lexical variation. The revelations that were uncovered through the method and analytical process, I would argue, has provided a template for the study of lexical variation for minority language speaking groups.

Lexical variation, as a subject for inquiry, has been largely ignored for numerous reasons outlined within this research (Durkin, 2012). Practical and theoretical implications concerning research that examined variation of lexis has been affected by various factors. This research has utilised methodology and theory that has merged interactional accounts of identity with observations and accounts of lexical variation. This approach has uncovered a rich and fine-grained account of identity as practised by individuals, but also detailed the use of lexical variation as an integral linguistic resource for identity. Influential background factors for

variation between individuals was also facilitated through this methodological approach. Future research that investigates minority groups, lexical variation, and how linguistic and, specifically, lexical variation has been used as a means for identity as social practice could help improve our understanding of language, including its variations within social groups, and the varieties that exist and are being used by individuals.

References

Acton, T. & Kenrick, D., 1984. Romani Rokkeripen To Divvus. *The English Romani Dialect and Its Contemporary Social, Educational and Linguistic Standing*. London: Romanstan Publications. (AK).

Acton, T., 1994. *Categorising Irish Travellers*. In eds. McCann et al. 36-53 approach. Discourse Studies, 7(4–5), pp.585–614.

Aitchison, J., 1997. *The Language Web. The Power and Problem of Words*. Cambridge: Cambridge University Press.

Albury, N.J., 2017. *The power of folk linguistic knowledge in language policy*. Language policy, 16(2), pp.209–228.

Austin, J. L. 1962. *How to do things with words*. Cambridge, MA: Harvard Univ. Press.

Ayla, J., and Vandelvelde, I. 2016. *Representing Rajasthani roots: Indian Gypsy identity and origins in documentary films.* Romani Studies, 2016, Vol.26(2), pp.151-173.

Bakker, P. & Mous, M. (Eds.), Mixed languages: 15 Case Studies in Language Intertwining (pp. 123-150). Amsterdam: IFOTT/Foris.

Bakker, P. 2002. An Early vocabulary of British Romani (1616): *A linguistic analysis. Romani studies*. Fifth Series, 12 (2): 75-101.

Bakker, P., and Matras, Y. 1997. Introduction. In: Matras, Y., Bakker, P., and Kyuchukov, H. eds. The typology and dialectology of Romani. Amsterdam: Benjamins. vii-xxx.

Bakker, P, and Matras, Y. 2013. Contact Languages. A Comprehensive Guide, ed. De Gruyter, Inc.

Baoill (Eds.), *Travellers and Their Language* (pp. 47-63). Belfast: Cló Ollscoil na Banríona.

Bataillard, P. 1884. *De l' Apparition des Bohemieus en Europe, Paris*. Available at: https://gallica.bnf.fr/ark:/12148/bpt6k62116166/f57.image.texteImage [Accessed: November 2018].

Beal, J., 2010. *An Introduction to Regional Englishes: Dialect Variation in England*. Edinburgh: Edinburgh University Press.

Belton, B.A., 2005. Questioning Gypsy Identity : Ethnic Narratives in Britain and America. Blue Ridge Summit: AltaMira Press.

Beníšek, M., 2020. The Historical Origins of Romani. In Matras, Y, & Tenser, A (eds) 2020, *The Palgrave Handbook of Romani Language and Linguistics*, Springer International Publishing AG, Cham.

Beníšek, M., 2006. "Ke kořenům slova rom." [On the roots of the word rom.] Romano džaniben, jevend. ppl9–28.

Binchy, A., 1993. *The status and functions of Shelta*. [Unpublished thesis] University of Oxford.

Binchy, A., 1994. Travellers' Language: A sociolinguistic Perspective. In eds. McCann et al. pp.134-54.

Binchy, A., 2002. Travellers' Use of Shelta. Travellers and their Language. eds. Kirk and Baoill, P.O. Queens Belfast University Belfast

Boorde, A. 1547. *The first boke of the introduction of knowledge*. Reprinted on F.J. Furnival (1907).

Boretzky, N. and Igla, B., 2004. *Kommentierter Dialektatlas des Romani*. Wiesbaden: Harrassowitz.

Elšík, V. and Beníšek, M., 2020. In: The Palgrave Handbook of Romani Language and Linguistics, edited by Yaron Matras, and Anton Tenser, Springer International Publishing AG, 2020.

Borrow, G.A., 19th C. Romani Luvo Lil: in Robert Dawson Special Collection. (B2).

Braber, N., 2015. Language perception in the East Midlands in England: Investigating East Midlands adolescents' perception of language variation in the UK. English Today, Vol.31.1., pp.16-26.

Braber, N., 2018 Pit Talk in the East Midlands. In: N. Braber and S. Jansen (eds.) *Sociolinguistics in England*. London: Palgrave Macmillan pp.243-274.

Braun, A., 2018. Unpublished analogue scale conversion software for Windows. almut.braun@york.ac.uk.

Bright, R., 19th C. Journal of the Gypsy Lore. New Series. Vol IX Part 3-4: in Robert Dawson Special Collection. (B1).

Bucholtz, M. and Hall, K., 2005. Identity and interaction: a sociocultural linguistic approach. Discourse Studies, 7(4–5), pp.585–614.

Bucholtz, M., 2012. *From Stance to Style. Gender, Interaction and Indexicality in Mexican Immigrangt Youth.* Chapter in: Jaffe, A., 2012. Stance sociolinguistic perspectives. New York . Oxford University Press.

Bucholtz, M., 1999. "Why be normal?": Language and identity practices in a community of nerd girls. In Language in society, 28(2), pp.203–223.

Butler, J. (1988). Performative acts and gender constitution: An essay in Phenomenology and feminist theory. Theatre Journal 40:519-31.

Butler, J., 2015. Gender Trouble: Feminism and the Subversion of Identity (1990). English studies in Canada, 41(4). p.15.

Butler, J., 1993. Critically Queer. GLQ, 1(1), pp.17–32.

Calso. 20th C. Journal of the Gypsy Lore Society, 3rd series, Volume III/VIII Part 3 and XXVIII parts 1-2. (V1).

Cash, A., 1977. *The Language of the Margins*. Journal of the Gypsy Lore Society 4th ser. 1.3.pp.177-80.

Chambers, J. K. and Trudgill, P., 1980. *Dialectology*. Cambridge Textbooks in Linguistics. Cambridge University Press.

Clark, L., & Trousdale, G. (2013). Using participant observation and social network analysis. In M. Krug & J. Schlüter (Eds.), Research Methods in Language Variation and Change (pp. 36-52). Cambridge: Cambridge University Press.

Cohen, A. P., (1985). *The Symbolic Construction of Community*. Chichester: Ellis Horwood.

Condon, L. et al., 2019. Engaging Gypsy, Roma, and Traveller Communities in Research: Maximizing Opportunities and Overcoming Challenges. Qualitative health research, 29(9). pp.1324–1333.

Coughlen, T. 2001. *Now shoon the Romani gillie: Traditional verse in the high and low speech of the Gypsies of Britain* . Cardiff: University of Wales Press.

Cressy, D., 2016. Trouble with Gypsies in Early Modern England, The Historical Journal, Vol. 59 (1). pp. 45-70.

Creswell, J.W. & Plano Clark, V. L., 2018. Designing and conducting mixed methods research Third edition..; International student., Los Angeles: SAGE.

Crofton, B & Smart, H., 1875. *The dialect of the English Gypsies*. 2nd Edition. London: Asher. (SC1)

Crofton, H. T. 1888. Early Annals of the Gypsies in England, Journal of the Gypsy Lore Society, Jul 1, Vol 1(1), p.5.

Crofton, H. T., 1907. Supplementary Annals of the Gypsies in England Before 1700, Journal of the Gypsy Lore Society, Vol. 1. p.31.

CSO (Central Statistics Office), 2011. [online]. [Accessed November 2018]. Available from: https://www.cso.ie/en/releasesandpublications/ep/p-cp8iter/p8iter/p8itd/

CSO (Central Statistics Office), 2016. [online]. [Accessed November 2018]. Available from: https://www.cso.ie/en/releasesandpublications/ep/p-cp8iter/p8iter/p8itd/

Dawson, R., 2002. The Dialect of Derbyshire's Traditional Travellers. Self-published. ISBN 1-903418-28-3. (DV2).

Dawson, R., 2009. An English to Romany/Scottish/Irish Traveller Dictionary. Self-published. ISBN 978-1-903418-66-6. (DV5).

Dawson, R., 2011. An Irish Traveller Pavee Cant (Gammon/Shelta) Dictionary. Self-published. ISBN 978-1-903418-78-9. (DV4).

Dawson, R., Special Collection. 20th C. University of Reading Special Collections and The Museum of English Rural Life. Collection reference D BD 1/8 – D BD 1/32. (DV3).

Dawson, R., Dawson Vocab 1959-69. Special Collection. University of Reading Special Collections and The Museum of English Rural Life. Collection reference D BD 1/8 – D BD 1/32. (DV1).

Dawson. R.M.,20th C. Words Collected in the North Country, particulary Yorkshire, West Riding: in Robert Dawson Special Collection. (D1).

Devlin, P., 2014. *Sociophonetic Variation, Orientation and Topic in County Durham*. PhD Thesis. York: University of York.

DeCarlo, C., 2017. The Byzantine Empire, edited by, Britannica: Educational Publishing. Rosen Publishing Group, 2017.

Dictionary of Scott's Language., 2021. [online]. [Accessed December 2020]. Available at https://dsl.ac.uk/

Diehl, D.K., 2019. Language and interaction: applying sociolinguistics to social network analysis. Qual Quant 53, pp.757–774

Dow, J.R., Niedzielski, N.A. & Preston, D.R., 2001. Folk Linguistics. The Journal of American folklore, 114(454), p.504.

Du Bois, J., 2007. The stance triangle. In *Stancetaking in discourse*. Ed Robert Englebretson, pp.139-182. Amsterdam: Benjamins.

Duffy, R., Duffy, Q., Duffy, V., McCreedy, T., Patel, J., Smith E. and Spencer, S., 2014. Patrin. Gypsy Heritage in the East Midlands. The Monitoring Group Pub.

Durkin, P., 2012. Variation in the lexicon: the 'Cinderella' of sociolinguistics? In English today, 28(4), pp.3–9.

East Anglian: No author. 1929. Anglo-Romani gleanings: from East Anglian Gypsies. Journal of the Gypsy Lore Society, Third Series, 8: pp.105,134. (EA).

Eckert, P., 1989. Jocks and Burnouts: Social Identity in the High School Eckert, P.New York: Teachers College Press.

Eckert, P., (1996) 'Vowels and nail polish: The emergence of linguistic style in the preadolescent heterosexual marketplace' In Cameron, D. & Kulick, D (eds.) The Language and Sexuality Reader, London: Routledge, pp. 189-195.

Eckert, P., 1997. Why ethnography? In U.-B. Kotsinas, A.-B. Stenstrom, & A.-M. Karlsson (Eds.), Ungdomsspråk i Norden (pp. 52-62). Stockholm: Stockholm University.

Eckert, P., 2000. *Linguistic Variation as Social Practice*. Oxford: Blackwell.

Eckert, P. 2012. Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation Annual Review of Anthropology, VOL 41; 41. pp.87-100

Eckert, P., 2018. *Meaning and Linguistic Variation*. Cambridge: Cambridge University Press.

Eckert, P., and McConnell-Ginet, S., 2003. *Language and gender*. Cambridge, Cambridge University Press.

Elšík, V., and Matras, Y., 2006. *Markedness and language change: The Romani sample*. Berlin: Mouton.

Kroonen, Guus., 2009. Etymological Dictionary of Proto-Germanic (Leiden Indo-European Etymological Dictionary) (English and Germanic Languages Edition) Bilingual Edition.

Evans, I.H.N., 1929a. Gleanings from English Gypsies. Journey of Gyspy Lore Society, New series, 8 (3). pp.140-142. (EV).

Farah I., (1997) Ethnography of Communication. In: Hornberger N.H., Corson D. (eds) *Encyclopedia of Language and Education. Encyclopedia of Language and Education* vol 8. Springer, Dordrecht.

Fishman, F. and García, O., (eds.) 2010. *Handbook of language and ethnic identity*. 2nd edn. Disciplinary and regional perspectives. New York: Oxford University Press.

Fox, Samuel. 1926. *The dialect of the Derbyshire Gypsies*. JGLS, 3rd series, 5, pp.62- 94. (F1).

Friedman, V. A. and Dankoff, R., 1991. The earliest known text in Balkan (Rumelian) Romani: a passage from Evliya Celebis SeyÇhat- nÇme. JGLS, fi fth series, 1: pp.1–20.

Gharibi, K.and Boers, F., 2017. Influential factors in incomplete acquisition and attrition of young heritage speakers' vocabulary knowledge. Language Acquisition, 02 January 2017, Vol.24(1), pp.52-69 . Psychology Press.

Gmelch, S., Langan, P., Gmelch, G., 1976. *Tinkers and Travellers*. McGill-Queen's University Press.

Grant, A. P., (1994). Shelta: The secret language of Irish Travellers viewed as a mixed language. In P.

Grellmann, H., (1783) *Die Zigeuner. Ein historischer Versuch*, Dessau

Griffiths, J. & Yates, D. E., 1934. Sanderson's vocabulary, part 4: Rommano-English Vocabulary. In: Journal of the Gypsy Lore Society, third series, 13. pp.59-88. (GY).

Griffiths, J. & Yates, D. E. 1934. Sanderson's vocabulary, part 4: Rommano-English Vocabulary. In: Journal of the Gypsy Lore Society, third series, 13. pp59-88. (GY).

Grosvenor, Lady Arthur., 1908. Whiter's Lingua Cinguariana. In, Journal of the Gypsy Lore. (G).

Halliday, M.A.K., 1976. American Anthropologist. New Series, Vol. 78, No. 3, pp. 570-584.

Hamp. No author., 1929. IV.-Anglo-Romani gleanings: (III) Chiefly from Hampshire. JGLS, 3rd Series, 115-139. (H).

Hamp. No Author. 1929. IV-Anglo-Romani gleanings: (III) from Hamspshire. Journal of the Gypsy Lore Society, 3rd series, 115-139. (H).

Hancock, H., 1975. *Problems in the creation of a standard dialect of Romanes Social Sciences* Research Council Working Papers No. 25. Austin Texas.

Hancock, H., 1984. Romani and Angloromani. In: Trudgill, P. *Language in the British Isles*. Cambridge: Cambridge University Press, pp. 367-383. 'Shelta and Polari', P. Language in the British Isles, pp. 367-383. Cambridge: Cambridge University Press.

Hancock, H., 1992. The Roots of Inequity: Romani Cultural Rights in their Historical and Social Context. *Immigrants and Minorities* Vol. 11 (1), pp.3-20.

Hock, H. et al., 2016. The Languages and Linguistics of South Asia : A Comprehensive Guide, edited by Hans., De Gruyter, Inc.

Holmes, J. & Meyerhoff, M., 1999. The Community of Practice: Theories and methodologies in language and gender research. In, Language in society. (it) 28(2), pp.173–183.

Holmes, J., Hazen, Kirk & ProQuest., 2014. *Research methods in sociolinguistics : a practical guide*. Hoboken: John Wiley & Sons Inc.

Howley, G., 2016. *The acquisition of Manchester dialect variants by adolescent Roma migrants*. Manchester: University of Salford.

Hudson, R. A., 1975. The meaning of questions. Language 51, pp.1–31.

Hymes, D., 1962. The ethnography of speaking. In, Anthropology and human behavior it 13(53), pp.11-74.

Hymes, D., 1964. *Language in Culture and Society: A Reader in Linguistics and Anthropology.* New York: Harper & Row.

Hymes, D. 1980. Language in Education: Ethnolinguistic Essays. In, *Language and Ethnography Series*. Washington, D.C.: Center for Applied Linguistics.

Hymes, D.H., (1975). Breakthrough into performance. In, D. Ben-Amos & K. Goldstein (Eds.), *Folklore: Performance and communication* it. (pp. 11–74). The Hague: Mouton.

Independent. 2000. [online]. [Accessed 20th November 2018]. Available from: https://www.independent.co.uk/news/uk/this-britain/irish-travellers-gain-legal-status-of-ethnic-minority-710768.html

Irvine's Vocabulary. No Author. Journal of the Gypsy Lore. 19th C. 3rd Series 4. 172-182: in *Robert Dawson Special Collection*. (IV) it.

Johnstone, B., and Marcello, W., 2010. Dell Hymes and the Ethnography of Communication. *The Sage Handbook of Sociolinguistics*. London: Sage.

Jones, L., 2020. Subverting transphobia and challenging ignorance: The interactive construction of resistant identity in a community of practice of transgender youth. In, *Journal of Language and Discrimination*. 4(2), 149–171.

Jones, L., 2012. *Dyke/Girl: language and identities in a lesbian group*. Houndmills, Basingstoke, Hampshire; New York: Palgrave Macmillan.

Jones, L., 2019. Discourses of transnormativity in vloggers' identity construction. In, International Journal of the Sociology of Language. 256(2), pp.85-10.

Kaufman, T., 1979. Review of W.R. Rishi, Multilingual Romani dictionary. In, Hancock (1979) pp. 131-144.

Keisling, S. F., 2004. Dude. American Speech 79 (3):pp. 281-305.

Keisling, S. F., 2012. Style as Stance. Stance as the Explanation for Patterns of Sociolinguistic Variation. Chapter in Jaffe, A., 2012. Stance sociolinguistic perspectives it. New York;: Oxford University Press.

Kerswill, P., Llamas, C. and Upton, C., 1999. The First SuRE Moves: Early Steps Towards a Large Dialect Project. In, *Leeds Studies in English* Vol.30, pp.257.

Kluge, F., 1901. Rotwelsch: *Quellen und Wortschatz der Gaunersprache und der verwandten Geheimsprachen*. Strasburg: Karl Trübner.

Kluyver, A., 1910. Un glossaire tsigane du seizième siècle. JGLS, new series, 4: pp.131–42.

Koptjevskaja-Tamm, M., 2000. Romani genitives in cross-linguistics perspectives. In, *Grammatical relations in Romani: The noun phrase*. (eds). Viktor Elsik and Yaron Matras, pp.123-149. Amsterdam: John Benjamins.

Kroskrity, P.V., 2009. Discourses of Endangerment: Ideology and Interest in Defence of Languages. In, Language in Society.it 38(4), pp.513–516.

Labov, W., (1966). The Social Stratification of English in New York.it Washington: Center for Applied Linguistics.

Labov, W., (1972a). Language in the inner city: Studies in the Black English vernacular it. Philadelphia: University of Pennsylvania Press.

Labov, W. (1972b). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Labov, W. 1963. The social motivation of sound change. Volume 19 – 3: 273-309.

Lakoff, Robin., (1975) *Language and a Women's place*. Oxford: Oxford University Press.

Lave, J. and Wenger, E., 1991. *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

Le Page, R. B. & Tabouret-Keller, A., 1985. *Acts of Identity: Creole-Based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.

Leeds Special Collection., 1979. Can you pooker Romany? Gipsy and Tinker cant of Britiain and Ireland. cassette tape. In: *Gypsy, Traveller and Roma*. Collections. BC MS Rom/Fraser/I/1/6 (LSP).

Legislation.gov.uk., 2010. [online]. [Accessed 14th January 2019]. Available from: https://www.legislation.gov.uk/ukpga/2010/15/contents

Leland, C. G., 1882. Shelta, the Tinkers' Talk. The Gypsies it. Boston: Houghton, Mifflin and Co. pp.354-372; reprinted as: The Tinkers'Talk. In, Journal of the Gypsy Lore Society 2nd ser. 1 (1907-08): 168-80.

Leland, C. G., and MacAlister, R. A., 1937. The Secret Languages of Ireland. In, Nature 139 it pp. 692–693.

Liverpool University special collection. Reference: Romani and Shelta notebooks – c.1885-1931. GLS C/3/2. Liverpool: Liverpool University.

Llamas, C., 1999. A New Methodology: Data Elicitation for Social and Regional Language Variation Studies. In, *Leeds Working Papers in Linguistics and Phonetics*. [Accessed September 2018]: http://www.leeds.ac.uk/arts/download/1322/llamas1999

Llamas, C. Watt, D., 2014. Scottish, English, British?: Innovations in Attitude Measurement: Innovations in Attitude Measurement. In, *Language and Linguistics Compass* Vol.8(11), pp.610-617.

Lucas, J., 20th C. *Journal of the Gypsy Lore. New Series*. Vol 8 Part 4. In Robert Dawson Special Collection. (L1).

Macalister, R. A. S., 1937. The Secret Language of Ireland. Manuscripts of the late John Sampson. Cambridge: Cambridge University Press; facsimile edition., Armagh: Craobh Rua Books, 1996.

MacGowan, A., 1996. The Winchester confessions 1615-1616. Depositions of travellers, Gypsies, fraudsters and makers of counterfeit documents, including a vocabulary of the Romany language. South Chailey (East Sussex): *Romany and Traveller Family History Society: in Manchester Corpus*. (MG).

Pauwels, A., (2016) Trends and patterns in language maintenance and shift. In, Language Maintenance and Shift (*Key Topics in Sociolinguistics* pp. 83-99). Cambridge: Cambridge University Press.

Manchester (Corpus) Romani Project. 2021. [online]. [Accessed 5 January 2018 - October 2021]. Available from: http://romani.humanities.manchester.ac.uk/ (MC).

Matras, Y. & Leggio, D. V., 1 Dec 2017, In: Romani Studies it. 27, 2, pp. 173-209 36 p., doi: https://doi.org/10.3828/rs.2017.10.

Matras, Y. & Tenser, A., 2020. *The Palgrave handbook of Romani language and linguistics* Cham, Switzerland: Palgrave Macmillan.

Matras, Y., 2003. Mixed languages: re-examining the structural prototype. In: Bakker, Peter & Matras, Yaron. (eds). *The Mixed Language Debate*. Berlin: Mouton de Gruyter. pp.151-175.

Matras, Y., 2002. *Romani: A Linguistic Introduction*. Cambridge: Cambridge University Press.

Matras, Y., 2010a. *The afterlife of a Language*. Edinburgh: Edinburgh University Press.

Matras, Y,. 2013. *Mapping the Romani dialects of Romania*. Romani Studies 5, Vol. 23, No. 2 (2013), pp.199–243.

Matras, Y., 2012, Dynamics of contact-induced Language Change. Leglise, I. & Chamoreau, C. (eds.). Berlin: de Gruyter, Walter GmbH & Co, pp. 1-28

Matras, Y., 2005. The classification of Romani dialects: A geographic-historical perspective. In, (eds.) Halwachs, D., Schrammel, Barbara, & Ambrosch, Gerd. *General and Applied Romani Linguistics*. Munich: Lincom Europa. pp. 7-26.

Matras, Y., 2010b. Contact, convergence and typology. In, Hickey, Raymond. (ed.) *Handbook of Language Contact*. Oxford: Blackwell. pp. 66-85.

Matras, Y., 2015. Why morphological borrowing is dispreferred. In: Gardani, F. et al. (eds.) *Borrowed Morphology*. Berlin: Mouton de Gruyter. pp.47-80.

Matras, Y, & Tenser, A (eds.), 2020, The Palgrave Handbook of Romani Language and Linguistics, Springer International Publishing AG, Cham. Available from: ProQuest Ebook Central. [25 November 2022].

Mayall, D., 1992. The making of British Gypsy Identities, c. 1500-1980. In, *Immigrants and Minorities*, Vol. 11 (1), pp.21-41.

Mayall, D., 1997. Egyptians and vagabonds: Representations of the gypsy in early modern official and rogue literature. In, *Immigrants and Minorities*, Vol. 16 (3), pp. 55-82.

Mayall, D., 2004. Gypsy identities 1500-2000: *From Egipcyans and Moon-men to the ethnic Romany*. London: Routledge.

Meakins, F., 2013. Contact Languages : A Comprehensive Guide, edited by Peter Bakker, and Yaron Matras, De Gruyter, Inc.

McColl, M. R., Barras, W. and Bonnici, L. M., 2014. *Lexical Variation and Attrition in the Scottish Fishing Communities*. Edinburgh: Edinburgh University Press.

McCooey-Heap, K. M., 2020. Putting a Sight on the Substrate, Yessir: *Lexical and syntactic items from the Manx Gaelic substratum in Manx English as resources for marking a Manx linguistic identity*. Doctoral Thesis. Birmingham: Edge Hill University.

McDonagh, M., (1994). Nomadism in Irish Travellers' identity. In, M. McCann, S. Ó Síocháin and J. Ruane (Eds.), *Irish Travellers: Culture and Ethnicity* (pp. 95-109). Belfast: Institute of Irish Studies, Queen's University Belfast.

Meli, G., 2016. From inflection to derivation: Outcomes of Early Romani genitive in Peidmontese Sinti. SKASE *Journal of Theoretical Linguistics* 13 (3): pp 29-45.

Meyer, K., 1891. On the Irish origin and age of Shelta. In, *Journal of the Gypsy Lore Society* 2,5: pp. 257-66.

Meyer, K. 1909. The secret languages of Ireland. In, *Journal of the Gypsy Lore Society*, New Series, 2: pp.241-6.

Miklosich, F. 1872–80. *Über die Mundarten und Wanderungen der Zigeuner Europas X–XII*. Vienna: Karl Gerold's Sohn.

Milroy, L. & Muysken, P., 1995. *One speaker, two languages : cross-disciplinary perspectives on code-switching*. Cambridge: Cambridge University Press.

Milroy, L., 1982. Social network and linguistic focusing. In, (ed.) *Sociolinguistic variation in speech communities*. Edited by: Romaine, S. pp.141–152. London, UK: Edward Arnold.

Milroy, L., 1980. *Language and social networks*. Oxford: Blackwell.

Milroy, L. 1987. *Language and social networks* 2nd ed. Oxford, , UK: Blackwell.

Modan, G., 2007. *Turf wars: Discourse, Diversity, and the Politics of Place*. Malden, MA: Blackwell.

Morpugo, M., 2006. *Mr Nobody's Eyes*. HarperCollins Publishers

Ni Shuinear, S., 2002. *The Curious Case of Shelta. Travellers and their Language*. (Ed.) Kirk and Baoill, P.O. Queens Belfast University. Belfast: Queens etc.

Norwood. No author. 19th C. *A vocabulary of words in Gypsie language or Romanish*. Norwood Notebooks: Scott MacFie collections: in Robert Dawson Special Collection (N1).

Ó hAodha, M. (2002)., Travellers' language: Some Irish language perspectives. In, (eds.) J. M. Kirk & D. P. Ó.

O` Baoill, D. P., 1994. Travellers' Cant: Language or Register?. In (eds.) McCann et al., pp155-69.

Ochs, E., 1992. Indexing gender. In, Alessandro Duranti and Charles Goodwin (eds.). Rethinking context: *Language as an interactive phenomenon*. pp. 335 – 57. Cambridge and New York: Cambridge University Press.

Office for National Statistics. 2011. [online]. [Accessed November 2018]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/ethnicity/articles/whatdoesthe2011censustellusaboutthecharacteristicsofgypsyoririshtravellersinenglandandwales/2014-01-21.

Oireachtas Report on the Recognition of Traveller Status [online]. [Accessed November 2018]. Available from: https://data.oireachtas.ie/ie/oireachtas/committee/dail/32/joint_committee_on_justice_and_equality/reports/2017/2017-01-26_report-on-the-recognition-of-traveller-ethnicity-january-2017_en.pdf

Okely, J., 1983. *The Traveller Gypsies*. Cambridge: Cambridge University Press.

Okely, J., 2014. Recycled (mis)representations: Gypsies, Travellers or Roma treated as objects, rarely subjects. *People, Place and Policy* Online, pp.65-85.

O'Reilly, K., 2005. *Ethnographic methods*. Abingdon, UK & New York, USA: Routledge.

Orton, H. & Dieth, E., 1962-1971. *The Survey of English Dialects*. Leeds: Arnold.

Oxford English Dictionary., 2020. [online]. [Accessed October 2020]. Available at https://www.oed.com/

Patton, M.Q., 2002. *Qualitative research and evaluation methods*. 3rd ed. Sage Publications; Thousand Oaks, CA.

Pauwels, A., 2016. *Language maintenance and shift*. Cambridge: Cambridge University Press.

Pichler, H., Wagner, S.E. and Hesson, A., 2018. Old-age language variation and change: Confronting variationist ageism: Old age language variation and change. In, editor *Language and Linguistics Compass*. Vol.12(6).

Ingo, P., 2003. *Phonology and Morphology of Creole Languages*. De Gruyter Inc.

Pott, A.,. 1844–1845. Die Zigeuner in Europa und Asien. *Ethnographischlinguistische Untersuchung vornehmlich ihrer Herkunft und Sprache*. Halle: Heynemann

Prieler, M. 2010. Othering, racial hierarchies and identity construction in Japanese television advertising. In, *International Journal of Cultural Studies*. 13(5): pp.511-529.

Rieder, M., 2018. *Irish Traveller Language: An Ethnographic and Folk-Linguistic Exploration* Palgrave Macmillan. (EM).

Rieder, M., 2015. *Irish Traveller English. In, Further Studies in the lesser-known varieties of English*. Schreier, D., Trudgill, P., Schneider, E.W., and Williams, J.P., pp. 70-96. Cambridge University Press. Cambridge.

Robert Dawson Special Collection. 1870-2001. Museum of Rural Life. University of Reading. Items D BD 1/15 – 1/32.

Roberts. No author. 1912. Robert's Vocabulary. In, Journal of the Gypsy Lore Society it new series, 5. pp.177-192: in Robert Dawson Special Collection. (R2).

Romlex lexical database. 2020. [online]. [Accessed November 2020]. Available ar http://romani.uni-graz.at/romlex/

Rüdiger, J. C. C., 1782 [reprint 1990]. Von der Sprache und Herkunft der Zigeuner aus Indien. In: Neuester Zuwachs der teutschen, fremden und allgemeinen Sprachkunde in eigenen Aufsätzen, 1. Stück. Leipzig. pp. 37–84. Hamburg: Buske.

Russell, A., 1915. Tom Taylor's Anglo-Romani vocabulary. In, *Journal of the Gypsy Lore Society*, new series, 8, 281,287. (R).

Russell, A., 1916. Bright's Anglo-Romani Vocabulary. In, *Journal of the Gypsy Lore Society*. New Series. 9, 165-185: in Robert Dawson Special Collection. (R).

Sampson, J., 1891. *Journal of the Gypsy Lore*. Vol 3 Issue 2: in Robert Dawson Special Collection. (S2).

Sampson, J., 1911. Jacob Bryant; Being an analysis of his Anglo-Romani vocabulary, with a discussion of the place and date of collection and an attempt to show that Bryant, not Rudiger, was the earliest discoverer of the Indian origin of the Gypsies. JGLS New Series, 4: 162-94. (S3).

Sampson, J., 1923. On the origin and early migration of the Gypsies. In, Journal of the Gypsy Lore Society, Third Series, 2 (4), 156-169.

Sampson, J., 1923. On the origin and early migration of the Gypsies. In, *Journal of the Gypsy Lore Society*. Third Series, 2 (4). pp.156-169.

Sampson, J., 1926a. *The Dialect of the Gypsies of Wales*. Oxford: The Clarendon Press.(FS).

Sampson, J., 1926b. The dialect of the Derbyshire Gypsies. Compiled in the parish of Smalley by the Rev. Samuel Fox 1832-1833. In, JGLS Third Series, 5(4): pp.62-94.

Sanderson, 19th C. Journal of the Gypsy Lore. 3rd Series. Vol XIII Part 3: in Robert Dawson Special Collection. (S1).

Sandow, R., (2020). The Anglo-Cornish dialect is 'a performance, a deliberate performance': Ideological orientation and patterns of lexical variation in a peripheral dialect. In, *English Today*, 36(3), pp.77-84.

Scala, A., 2020. Romani Lexicon. Matras, Y, & Tenser, A (eds.), The Palgrave Handbook of Romani Language and Linguistics: Springer International Publishing AG, Cham.

Schreier, D., Trudgill, P., Edgar W. Schneider and Williams J. P., 2015. (eds.): The Lesser Known Varieties of English: An Introduction. Cambridge University Press. Cambridge.

Scotland Census., 2011. [online]. [Accessed November 2018]. Available from: https://www.gov.scot/publications/gypsy-travellers-scotland-comprehensive-analysis-2011-census/pages/3/

Secret Portsmouth., 2016. [online]. [Accessed December 2020]. Available at https://languagelearnersjournal.com/2018/01/22/the-ultimate-guide-to-pompey-slang/

Silverstein, M., 1976a. Shifter, linguistic categories, and cultural descriptions. Meaning in anthropology. (eds.) Keith H. Basso and Henry A. Selby, pp.11-56. Albuquerque: University of New Mexico Press.

Silverstein, M., 1985a. Language and the culture of gender: at the intersection of structure, usage, and ideology. In, Semiotic Mediation: Sociocultural and Psychological Perspectives (E. Mertz and R. Parmentier, (eds.)), 219-259. Orlando: Academic Press.

Silverstein, M., 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication*. 23(3–4), pp.193–229.

Sims, R. R., 2002. *Managing organizational behavior*. Westport, CT: Quorum Books.

Skinner, J. (Ed.).,2012. *The Interview: An Ethnographic Approach*. Routledge.

Smart, H. & Crofton, B., 1875. *The dialect of the English Gypsies*. 2nd ed., Revised and Greatly Enlarged. London: Asher.

Somerset Live. 2021. [online]. [Accessed December 2020]. Available at https://www.somersetlive.co.uk/news/somerset-news/romany-words-english-language-everyday-4353188

Suek, L.A., 2014. The Influence of Language Use and Language Attitude on the Maintenance of Community Languages Spoken by Migrant Students. In, (ed) Englisia, 1(2), pp.Englisia, 2014–05-01, Vol.1 (2).

Taylor, B., 2014. Another Darkness, Another Dawn : A History of Gypsies, Roma and Travellers. Reaktion Books, Limited.

Taylor, T., 1915. Word list. In Russell, A. 1915. Tom Taylor's Anglo-Romani vocabulary. In, *Journal of the Gypsy Lore Society*, new series, 8, 281,287. (T1).

Tedlock, B., 1991. From Participant Observation to the Observation of Participation: The Emergence of Narrative Ethnography. In, ed. *Journal of anthropological research*, 47(1), pp.69-94.

Trudgill, P., 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. In, ed. Language and Society 1:179-95.

Turner, R., 1924. *Indo-Aryan etymological notes*. BSOS, iii, 2 399-404.

Turner, Ralph., 1926.The Position of Romani in Indo-Aryan. In, ed. Journal of the Gypsy Lore Society: 3rd series, Vol. IV, No. 4, pp.146-189. Gypsy Lore Society: Liverpool.

Underwood, G. N., 1988. Accent and Identity. In, Thomas, A.R., (ed.) *Methods in Dialectology.* Clevedon: Multilingual Matters. pp.406-427.

Various Sources 1. 20th C. Journal of the Gypsy Lore Society, 3rd series, Volume XXVII Part 3-4 and XXVIII parts 1-2. (V2).

Various Sources 2. 20th C. Journal of the Gypsy Lore. 3rd Series. Vol II Part 4: in Robert Dawson Special Collection. (V1).

Velázquez, I., 2013.Mother's social network and family language maintenance. In, *Journal of Multilingual and Multicultural Development* Vol.34(2), pp.189-202.

Wagner, S.E., 2012. Age Grading in Sociolinguistic Theory. In, Language and linguistics compass, 6(6), pp.371–382.

Walters. R., 2002. The Trivialisation of the Cant. Travellers and their Language. Eds. Kirk and Baoill, P.O. Queens Belfast University. Belfast.

Wardhaugh, R. & Fuller, J. M., 2021. *An introduction to sociolinguistics.* Eighth edition. Chichester: Wiley Blackwell.

Way. 20th C. *Journal of the Gypsy Lore*. 3rd Series. Vol 34 Part 3-4: in Robert Dawson Special Collection. (W1).

Wentzel, T.W. and Cherenkov, L.N., 1976. Dialekti tsiganskogo jazika. Jazika Aziji I Afrika 1:15-26. Moscow.

Weyers, J.R., 2014. The "Tuteo" of Rocha, Uruguay: A Study of Pride and Language Maintenance. Hispania, 2014, Vol.97(3), pp.382-395.

White, H.C., 2008. *Identity and Control: How Social Formations Emerge*. Princeton, Princeton University Press.

Winstedt, E., 1913. Early British Gypsies. In, *Journal of the Gypsy Lore Society*, 1913, Vol. 7, p.5.

Winstedt, E. O. 1948. Anglo- Romani gleanings from the northern counties. *JGLS*. third series, 27: 83–110. (W2).

Winstedt, E. O. 19th C. Anglo-Romani gleanings from the Northern Counties. In, *Journal of the Gypsy Lore Society*, 3rd series, 28. 50-61. (W3).

Wolfram, W. & Schilling-Estes, N., 2005. American English: dialects and variation 2nd ed., Oxford: Blackwell.

Wolfram, W., 2000. Reconstructing the history of AAVE: new data on an old theme. In, Berkeley Linguistics Society, 26, pp.333-48.

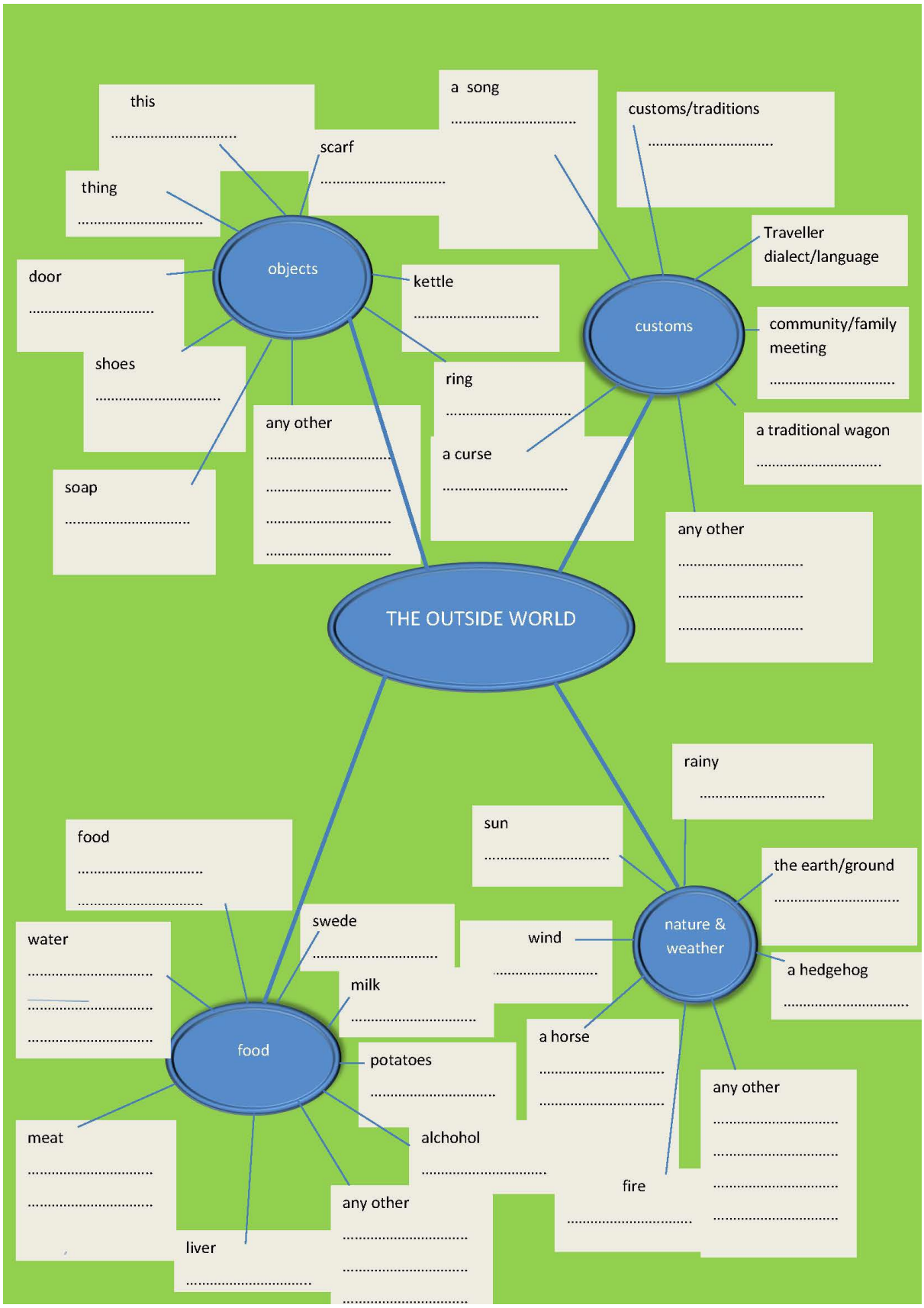
Woolner, A. C. 1928. Asoka and the Gypsies. *JGLS* 3d S. 7: pp.108-111.

Zhang, D., 2009. Language attitudes and heritage language maintenance among Chinese immigrant families in the USA. Language, culture and curriculum, 22(2): pp.77–93.

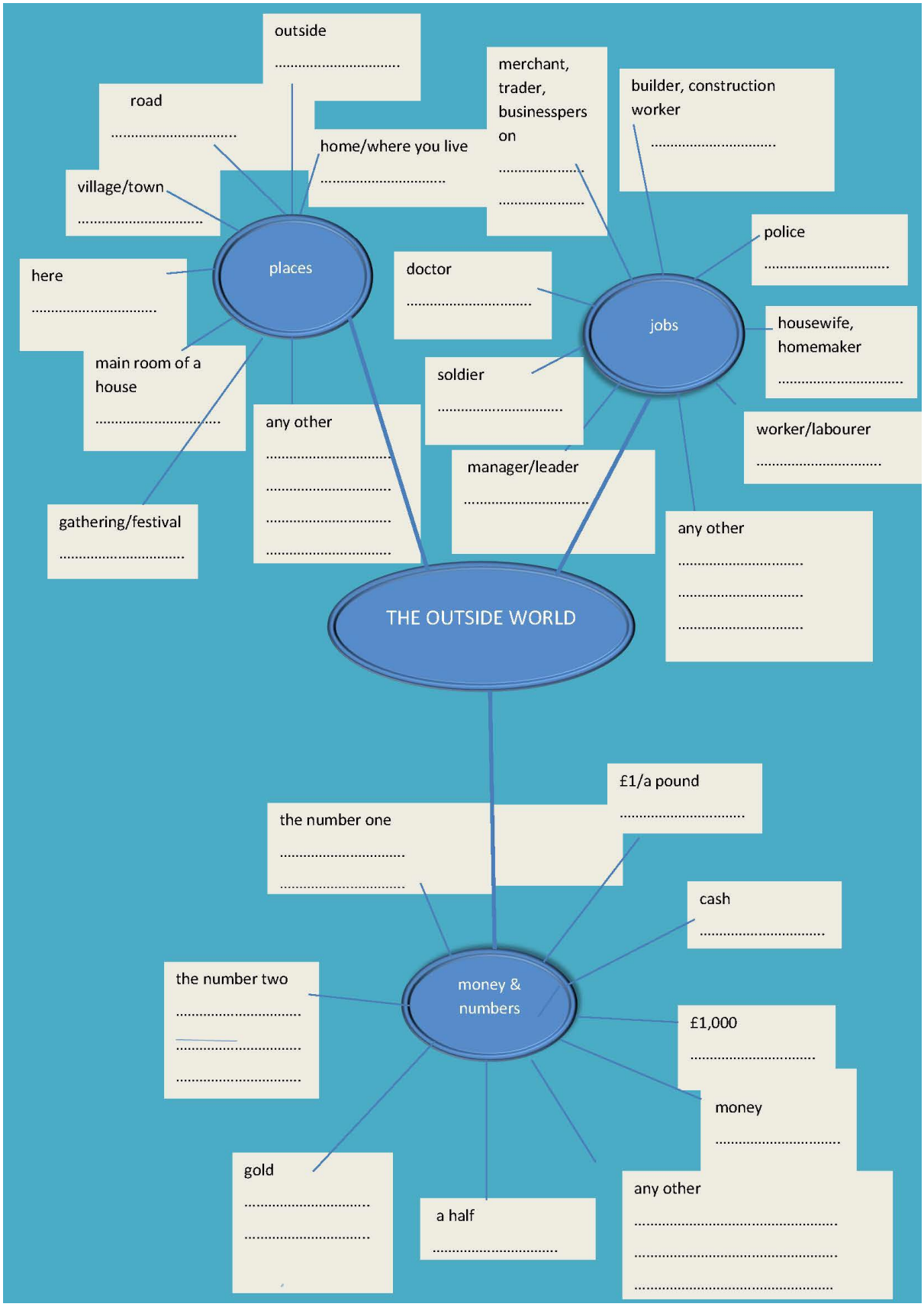
Zhang, D., 2012. Co-Ethnic Network, Social Class, and Heritage Language Maintenance Among Chinese Immigrant Families. In, Journal of language, identity, and education, 11(3), pp.200–223.

Appendix 1 Sense Relation Networks

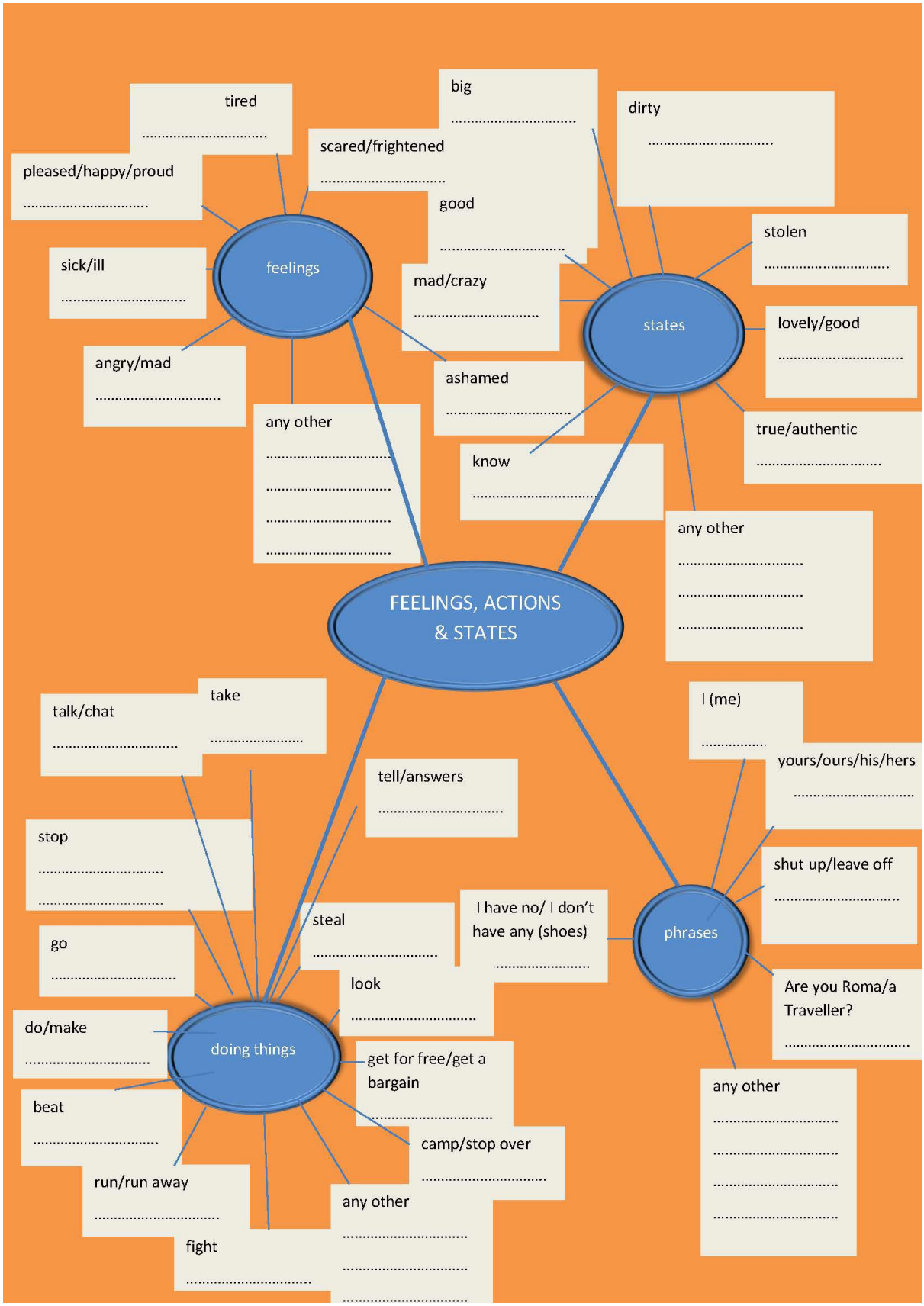
Sense Relation Network (SRN) The Outside World 1



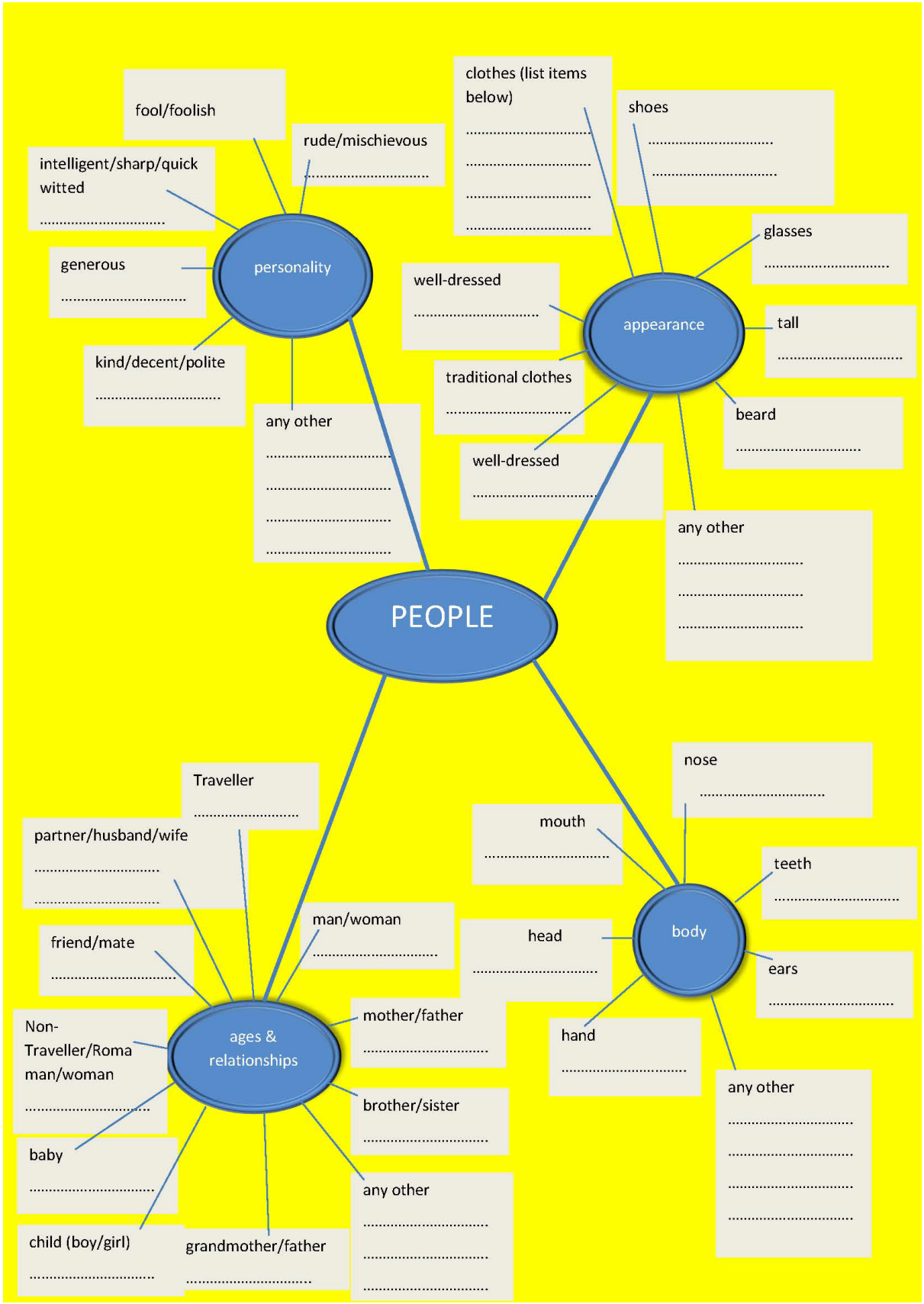
Sense Relation Network (SRN) The Outside World 2



Sense Relation Network (SRN) Feelings, Actions and States



Sense Relation Network (SRN) People



Appendix 2 Identity Questionnaire (ID Questionnaire)

Identity Questionnaire

Your Language and Identity

What accent would you say you had, and how would you describe it?

…….

Can you recognise the accent or dialect of your family and friends within the Traveller or Roma community?

…….

Can you recognise the accent or dialect of your local area?

…….

Do you think older and younger people talk the same amongst your family and friends and others within the Traveller or Roma community?

…….

Do you change the change the way you talk depending on the situation? If so, in what situation, and why?

…….

…….

…….

Do you think there's a difference between how males and females speak amongst your family, friends and others within the Traveller or Roma community? If so, in what way and can you give any examples of these difference?

…….

…….

…….

Are there any differences between the accent and dialect of your family, friends and others within the Traveller or Roma community and other people from the local area? If so, in what way and can you give any examples of these difference?

…….

…….

…….

Is there a difference between older and younger people speak within the Traveller or Roma community? Can you describe some of these differences?

…….

…….

……

Are there any differences between the accent or dialect of your Traveller or Roma community and other Traveller or Roma communities from the surrounding area? If so, can you describe any differences or give any examples?

……

……

……

How would you describe your identity either regionally, nationally, ethnically, or as a community?

……

……

……

Your Area

If you were watching a regional news programme, what places would you expect to hear news from ?

……

What local areas are important to you and your family and friends?

……

What image or description of your local area would you give to someone who didn't know it?

……

If you wanted a day out in your local area, where would you go and what would you do?

……

What do you consider the local football derby to be?

……

Outside of work, what activities do you consider common in your area?

……

What do you consider the best and worst things are about growing up and living in your area?

……

Have you ever seen your local area on a national T.V programme (e.g. a documentary)? If so, how was it portrayed?

……

If an outsider was complaining about your local area, would you defend it even if you agreed with what s/he was saying? Why/why not?

……

……

How many friends, relations and work/school/college mates do you have in the neighbourhood (not more than 10 minutes away) who you regularly see?

…….

Can you estimate how many close friends you have?

…….

How many of these do you consider as part of the Gypsy and Traveller community?

…….

How many do you consider being from outside the Gypsy and Traveller community?

1 = all friends are of the same ethnicity/community/group

2 = up to 20% of a different ethnicity/community/group

3 = up to 40% of a different ethnicity/community/group

4 = up to 60% of a different ethnicity/community/group

5 = up to 80% of a different ethnicity/community/group

Your community

Outside of work, what activities do you consider common amongst members of the Traveller or Roma community?

…….

…….

If an outsider was complaining about the Traveller or Roma community, would you defend it even if you agreed with what s/he was saying? Why/why not?

…….

…….

What do you consider the best parts about growing up and living as a member of the Traveller or Roma community?

…….

…….

Do you think there are any disadvantages growing up and living as a member of the Traveller or Roma community? If so, can you explain or give examples?

…….

…….

If you were to hear about Traveller or Roma community in the local or national news, what kind of stories would you expect to hear?

……

……

What regional areas are important to you and the Traveller or Roma community members?

……

……

Can you name any common destinations for members of the Traveller or Roma community within and outside of the local region?

……

……

What image or description of the Traveller or Roma community would you give to someone who didn't know about it?

……

……

Outside of work, what activities are common or popular amongst members of the Traveller or Roma community?

……

……

Are there any differences between the Traveller or Roma community groups within your region? Are there any differences between the Traveller of Roma community and other groups outside of this region?

……

……

What do you consider the benefits and drawback of life in the U.K as part of the Traveller or Roma Community?

……

……

Have you ever seen the Traveller or Roma community portrayed on T.V.(e.g. a documentary) ? If so, how as it portrayed?

……

……

What practices (e.g. superstitions, beliefs, customs) do you consider as being part of the Traveller or Roma community? Can you describe some/any of these?

……

……

……

……

Are there any customs or practices that older generations/relatives practiced (e.g. superstitions, beliefs, story telling, songs)? Can you describe any of these?

……

……

……

……

What percentage of your neighbourhood households are members of the Traveller or Roma community?

1 = all neighbours are of the same ethnicity/community/group

2 = up to 20% of a different ethnicity/community/group

3 = up to 40% of a different ethnicity/community/group

4 = up to 60% of a different ethnicity/community/group

5 = up to 80% of a different ethnicity/community/group

How many households do you have strong ties with in your neighbourhood/community

……

How many of these households are members of the Traveller or Roma community?

……

How many households do you have strong ties with outside of your neighbourhood/nearby community? How many of these households are members of the Traveller or Roma community?

……

Do you participate in daily activities/go to work/attend school with others from your neighbourhood? YES/NO

Are they members of the Traveller or Roma community? YES/NO

Are these people all male,all female or mixed?

……

How often do you associate with your work colleagues outside of work?

daily/weekly/fortnightly/monthly/less than this

Appendix 3 Identity Score Index Questionnaire (ID Score)

Identity Score Index Questionnaire

1. If you were on holiday and saw someone you had never seen before but thought they came from the Traveller or Roma community (e.g. you over heard their accent and recognised it, they were discussing Traveller or Roma community related topics etc.) would you:
    a. Feel compelled to go and ask where they were from and strike up a relationship
    b. Feel you had something in common but not do anything about it
    c. Not feel any differently than you would towards any other stranger
2. If you were on holiday and saw someone you had never seen before but thought they came from your local area (e.g. you over heard their accent and recognised it, they were wearing a local football shirt etc.) would you:
    a. Feel compelled to go and ask where they were from and strike up a relationship
    b. Feel you had something on common but not do anything about it
    c. Not feel any differently than you would towards any other stranger
3. Would you say you feel close to and feel you have something in common with people from the Traveller or Roma community in general (that is people you don't know personally), or would you say you do not feel any closer to them than to people from somewhere else?
    a. Feel closer to people from local area
    b. Don't feel any closer to people from local area than to other people
    c. Don't know, can't say
4. Would you say you feel close to and feel you have something in common with people from your local area in general (that is people you don't know personally), or would you say you do not feel any closer to them than to people from somewhere else?
    a. Feel closer to people from local area

b. Don't feel any closer to people from local area than to other people

c. Don't know, can't say

5. If you were the manager of a company which was recruiting people and two equally qualified and experienced people applied for the position, but one had raised within the traveller community and the other was from outside the community would you choose:

    a. The person from the Traveller or Roma community

    b. The person from outside

    c. Don't know, wouldn't matter

6. If you were the manager of a company which was recruiting people and two equally qualified and experienced people applied for the position, but one had been born and educated in your local area and the other had been born and educated somewhere else would you choose:

    a. The person from your area

    b. The person from somewhere else

    c. Don't know, wouldn't matter

7. Would you prefer your child's school teacher to be:

    a. A person from within the Traveller or Roma community familiar and speaks with a community dialect and/or accent

    b. A person who spoke 'standard' English with a 'standard' accent

    c. It wouldn't matter who they were

8. Would you prefer your child's school teacher to be:

    a. A local person with a local accent

    b. A person who spoke 'standard' English with a 'standard' accent

    c. It wouldn't matter what accent the had

9. If you were voting in a local election, would the fact that a candidate was a member of the Traveller or Roma community persuade you to vote for them?

a. Yes it would

b. No it wouldn't

c. Don't know

10. If you were voting in a local election, would the fact that a candidate was a local from your person persuade you to vote for them?

    a. Yes it would

    b. No it wouldn't

    c. Don't know

11. If you wanted to leave something to a charitable organisation would you choose:

    a. A charity connected to the Traveller or Roma community

    b. A national/international one

    c. Don't know, depends on the cause

12. If you wanted to leave something to a charitable organisation would you choose:

    a. A local one

    b. A national/international one

    c. Don't know, depends on the cause

13. If there was a programme on T.V. about your local area which clashed with your favourite programme and you could only watch one, would you:

    a. Watch it and miss your favourite programme

    b. Watch your favourite programme and miss the other one (but wish you hadn't)

    c. Watch your favourite programme and miss the other (but not mind)

14. If there was a programme on T.V. about the Traveller or Roma community which clashed with your favourite programme and you could only watch one, would you:

    a. Watch it and miss your favourite programme

    b. Watch your favourite programme and miss the other one (but wish you hadn't)

c. Watch your favourite programme and miss the other (but not mind)

Appendix 4 Analogue Scales

Visual Analogue Scale

On the line below, put a cross on the line that best represents your identity

British

agree _____ **disagree**

English

agree _____ **disagree**

East Midlander

agree _____ **disagree**

Roma

agree _____ **disagree**

Irish Traveller

agree _____ **disagree**

Welsh Traveller

agree _____ **disagree**

Irish

agree _____ **disagree**

Welsh

agree _____ **disagree**

Scottish

agree _____ **disagree**

British Roma

agree **disagree**

Appendix 5 Participant Information Sheet

NOTTINGHAM TRENT UNIVERSITY

Participant Information Sheet

Lesser Known Variety of English and Cultural Identity: Dialect and sociolinguistic variation within the Traveller and Roma communities of the East Midlands

Invitation

You are invited to take part in research carried out as part of a Doctoral research study at Nottingham Trent University. Before deciding to participate in the study, it is important that the purpose of the research has been explained and you have opportunities to ask any questions. Please read the following information and discuss it with others if you wish. If you have any questions, please feel free to ask the researcher or research supervisor directly, or contact them at a later date (contact details are at the end of this information sheet).

Purpose of research

I am interested in regional accent and dialect in the East Midlands and would like to find out more about how accent and dialect vary across the areas of the East Midlands. Traveller and Roma culture has often been overlooked when considering local and regional variation of accent and dialect and I would like to collect examples of language use and knowledge from people with experience and a background from the communities and individuals that identify and consider themselves Roma or Traveller. To do this I would like to interview and learn more about people's experience and knowledge of their community, culture and heritage and help create a more detailed and informed and representative picture of a regional and cultural heritage.

Through this project, I aim to examine the everyday speech of individuals and community groups across the region of the East Midlands using data collected from within the Traveller or Roma communities from interviews and information sheets to better understand and represent regional language variation within the East Midlands. In order to do this, I am going to:

- Compile data on language variation from individuals using language data collection sheets
- Collect interview recordings to supplement data from material representing language and culture from the regions of the East Midlands
- Create a written record of audio recordings from this sample with accompanying notes, for dialect and sociolinguistic purposes.

Why have you been invited to participate?

I would like to include people who identify themselves as being from Traveller or Roma communities and those who consider themselves as having Traveller or Roma culture as part of their heritage. Your participation in this study will contribute to a greater understanding of cultural heritage and how cultural identity and language may be maintained and transmitted for next generations.

What will happen if you take part?

I would like to interview you either with a friend or in a small group (3 or 4 people) after you have answered a few brief questions on a form given a week before the interview. I will also ask you to complete a short survey during the interview. If after the interview you are willing, I may contact you again to ask a few follow-up questions to explore your ideas in more depth. In total I am hoping to talk to about fifty to sixty people and hope that you agree to be one of them. If you are willing to talk to me about your knowledge and use of Traveller or Roma language, and Traveller or Roma Identity, I will ask you to share your thoughts with me in an informal pair or group discussion which may last up to one hour. This chat will take place in an area of your choosing (for example your/friends home, local community centre etc.).

You are free to withdraw or refuse to answer any questions at any time during the study and your participation is voluntary at all times. If you wish to withdraw at any point during your participation please let the researcher or research supervisor know that this is your wish.

How will the information you share be used?

Your data will only be used for research purposes. If you decide to take part I will ask your permission to record the interview which will be downloaded on to a password protected secured computer and accessible to the researcher and research supervisor only. Data to be shared or published will be anonymised. This means any personal data (such as your name) and all identifiable information will be removed so no one (apart from me) will know who said what. All information you share with me will be stored on password protected computers and backed up on Nottingham Trent University's own protected storage service, in accordance with the university's research regulations and also GDPR requirements 2018.

The anonymised interview and survey results will be moved to Nottingham Trent Universities data archive so that other researchers can access and use it. It will stay on this archive for 10 years.

You will always have the right to request access to your personal data and request correction and removal of personal data.

You also have the right to lodge any complaint with the Information Commissioner's Office (ICO). Please ask the researcher or research supervisor for information about this procedure.

Who has approved the project?

This research has been approved by the School of Arts and Humanities at Nottingham Trent University.

What do you have to do?

We will ask each volunteer to sign a consent form before any recording is made, which states that they are happy for an audio and/or written recording of their interview and survey responses to be held by the University for research purposes, and for relevant parts to be reproduced within publications and/or presentations. Any volunteer is free to withdraw from the study at any time, and without having to give a reason.

It is planned that this work will result in the publication of academic articles and the presentation of papers at national and international conferences.

If you have any questions at any time, please feel free to use the following contact details:

Peter Lee
Post Graduate Research
School of Arts and Humanities
Clifton Campus
Nottingham Trent University
Nottingham NG11 8NS
peter.lee@ntu.ac.uk

Dr Natalie Braber
PhD Supervisor
School of Arts and Humanities
Clifton Campus
Nottingham Trent University
Nottingham NG11 8NS
natalie.braber@ntu.ac.uk
0115 848 3011

Thank you for considering taking part in this research study.

Appendix 6 Consent form

NOTTINGHAM TRENT UNIVERSITY

Research Consent Form

Lesser Known Variety of English and Cultural Identity: Dialect and sociolinguistic variation within the Traveller and Roma communities of the East Midlands

I understand that this project will be carried out in accordance with the Research Ethics Codes of Practice of Nottingham Trent University, which can be viewed at:

https://www.ntu.ac.uk/__data/assets/pdf_file/0022/204727/code-of-practice-2018.pdf

1. I confirm that I have read or have had read to me the participant information sheet dated [ ] for the above study and have had the chance to ask questions and have had any questions answered to my satisfaction.	Yes ☐ No ☐
2. I understand that my involvement is voluntary and that I can refuse to answer questions and am free to withdraw at any point without giving a reason.	Yes ☐ No ☐
3. I understand that taking part in the study involves an audio-recorded interview and survey questionnaires completed by the researcher and participant.	Yes ☐ No ☐
4. I understand that my responses may be used in the researcher's dissertation, and research papers. When parts of my responses are used the researcher will take all possible actions to anonymise responses by removing any direct (e.g. names) and indirect identifiers (e.g. place names, events). All names will be removed from published transcripts.	Yes ☐ No ☐
5. I understand that personal information collected about me that can identify me, such as my name or where I live, will be restricted to the researcher and the research supervisors.	Yes ☐ No ☐
6. I give permission for interview and survey data that I provide to be deposited in NTU's data archive so that it can be used for future research and learning. The researcher will take all possible actions to anonymise the data by having direct (e.g. names) and indirect identifiers (e.g. place names, events) removed. The data will be accessible to other researchers with legitimate research interests to access and use for relevant research purposes upon request for a period of 10 years.	Yes ☐ No ☐
7. I accept that the researcher will keep the audio recording of my interview and survey documentation and these will only be available to the researcher and supervisor for the purposes of this and any related study. All recordings and documents will be transferred to a password protected computer and backed up on Nottingham Trent Universities central storage service and all other paper and digital copies will be erased or destroyed, following the GDPR 2018 and also Nottingham Trent University's data management policy and document retention schedule.	Yes ☐ No ☐

Name [PRINT] …… Signature …… Date……
(participant)

Name [PRINT] …… Signature …… Date……
(witness)

I have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

Name [PRINT] …… Signature …… Date……
(researcher)

Study contact details for further information

Peter Lee
Post Graduate Research
School of Arts and Humanities
Clifton Campus
Nottingham Trent University
Nottingham NG11 8NS
peter.lee@ntu.ac.uk

Dr Natalie Braber
PhD Supervisor
School of Arts and Humanities
Clifton Campus
Nottingham Trent University
Nottingham NG11 8NS
natalie.braber@ntu.ac.uk
0115 848 3011

Appendix 7

Data Management Plan

Lesser Known Variety of English and Cultural Identity: Dialect and sociolinguistic variation within the Traveller communities of the East Midlands

Data Collection

What data will you collect or create?

- This study will collect/ create the following data:
- Biographical data in plain text
- Interview data will be collected in the form of audio files
- Qualitative and quantitative survey data including plain text,point scales and spread sheet data
- Visual analogue scales in the form of point scales and corresponding image files

How will the data be collected or created?

- Data will be collected via recorded interviews and paper-based surveys.
- Survey results collected on paper will be digitised at the earliest opportunity to aid analysis, improve efficiency and safeguard the data.
- Visual analogue point scale results will be collected via a password protected laptop.
- Survey results will be tranfered from paper to digitial file and stored on a password protected laptop and NTU storage area. The paper copies will then be destroyed.
- Visual analogue point scale results will be colected via a password protected laptop and stored on NTU's digital storage area.

Documentation and Metadata

What documentation and metadata will accompany the data?

- The following study and file level documentation will be provided:
- File metadata will include codes that will show the survey results associated with the interview recordings.
- Details such as the transcription coding of the interviews, dates of when the interviews take place, file naming conventions and structure will be clearly documented.
- Procedural information regarding the method of collection and interpretations of the results and transcriptions will accompany the data in the form of a readme.txt file to allow future users to navigate through and understand the data.
- Upon completion of the study, a detailed metadata record of the finalised dataset will be created, using the
- DataCite Schema and added to NTU IRep to aid the discoverability of the data. This record will clearly link to any associated publications/ outputs.

Ethics and Legal Compliance

How will you manage any ethical issues?

- All data management procedures will be approved by the ethics board of Nottingham Trent University and any changes made during the process of the data collection and storage; further approval will be sought.
- The study will be fully GDPR compliant. The informed consent of each participant will be obtained. A copy of the participant information sheet and a blank copy of the informed consent form will be retained. Digital copies of the signed informed consent forms will be kept in an encrypted file. Participants will have the right to withdraw at any stage of the process and all of their data will be securely disposed of.
- Any personal and sensitive data will be carefully safeguarded (see Storage and Back Up section). In addition, audio interview data will be collected using a secure, pin protected device and downloaded to a password encrypted laptop as soon as possible after each interview whereupon the data on the recording device will be erased. Paper-based surveys containing personal data will be kept securely in locked filing systems until they are digitised, after which they will be securely disposed of in line with NTU policy.

- Personal biographical data will be used by the researcher to enable cross referencing of interview and survey data. This will enable the researcher to identify participants that may contacted for further data collection depending on the quality of the first round of data collection. Upon completion of the research project, the biographical data will have direct and indirect identifiers removed. The biographical data will only be available to the researcher and supervisors until the completion of the research project. Upon completion of the project, this biographical data will be stored on a password protected laptop for corresponding research purposes by the researcher only.
- Upon completion of the project, the anonymised biographical data will be archived with reuse of this data available upon request to members of the research community with relevant research enquiries.
- Survey and audio data will be only available by the researcher and the supervisors during the course of the project. Excerpts and quotations used in publications will have direct and indirect identifiers removed. Upon completion of the research project, all identifiers, direct and indirect, will be removed and made available to those researchers with legitimate and relevant research enquiries via the NTU Data Archive. The intention of the researcher is to make all data as openly available as possible without risk of disclosure. The risk of disclosure will be reassessed upon anonymization to assess more accurately for risk of disclosure when finalizing datasets for
- archiving given the relatively small number of participants.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

- The data collected will be academic and subject to academic conventions of use and resuse (i.e., appropriate and conventional forms of referencing) and follow licencing CC-BY-NC 4.0

Storage and Backup

How will the data be stored and backed up during the research?

- Data will be saved to designated folders on NTU's centralised storage service for active research data. This environment has a shared-nothing, distributed architecture, thus guaranteeing data redundancy and integrity.
- Access to this data storage will be restricted to those involved in the project. Small amounts of working data will be stored temporarily on an encrypted, password protected laptop, then saved to the networked storage facility.

How will you manage access and security?

- During the duration of the project, only the researcher and project supervisors will be able to access the data via NTU’s centralised storage service for active research data. As previously described, small amounts of working data will only be stored in analogue form and/ or on mobile devices for short periods of time in order to minimise risks of data loss and theft.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

- The researcher envisages research into accent and dialect, culture and identity, regional culture, ethnolinguistics and language variation over time would have interest in the use and access to this set of data.
- Interview, survey and anonymised biographical data will be retained and preserved for future reuse.
- All direct and indirect identifiers within the data will be erased. The anonymised data will be made available via NTU's Data Archive and accessible to those with legitimate and relavent research questions.

What is the long-term preservation plan for the dataset?

- The long-term preservation plan will mean data will need to be anonymised and processed with all indirect and direct identifiers removed. The files will be modified and amalgamated for searchability and attainability and transferred to the NTU Data Archive. Data will be preserved for 10 years.

Data Sharing

How will you share the data?

- Data will be shared upon completion of the research project. Data will be deposited in the NTU Data Archive. Upon deposit, a DOI (Digital Object Identifier) will be minted; this unique identifier can be used to cite the data in publications. An accompanying metadata record, using the DataCite Schema, will also be created in NTU’s IRep so that people can discover the data. This record describes the data, links to the thesis record and details the data access arrangements (detailed below).
- Use of the data will require users to follow appropriate academic conventions in terms of referencing and citation and as previously mentioned a CC-BY-NC 4.0 license will be applied to the data.

Are any restrictions on data sharing required?

- Restrictions will involve limiting access to the data to those that are legitimate researchers such as those with official affiliations and limited to those with relevant research enquiries to the data set. NTU Data Archive is a mediated service, therefore any requests to access the data will be submitted by email and permission from either the researcher or the supervisor will be sought in the first instance before access to the data is provided. This is to control for risk of disclosure given the relatively small number of participants, and so permitting only those with legitimate research interests relevant to the original purposes for the data collection. This will limit any possible risk of illegitimate use and assure participants that data collected will be limited to linguistics and associate research purposes.

Responsibilities and Resources

Who will be responsible for data management?

- The project researcher, Peter Lee, will take the lead role in supervising and maintaining and revising, managing changes during the research project period. These responsibilties include data capture, storage and backup, archiving, anonymising in preparation for archiving and ensuring these comply with the legal obligations outlined in the GDPR.

What resources will you require to deliver your plan?

- Resources needed for implementation of this plan will require consultation with appropriate staff within Nottingham Trent University such as project supervisors and data mangement advisory staff. Use and access to NTU's encrypted data repository and use of encripted data file transfer service will also be required. These services are available upon request to fully enrolled postgraduate researchers and PhD candidates who are handling personal and sensitive research data. A password secured recording device and laptop will also be provided by the institution and researcher respectively.

Data Protection Impact Assessment

What do you require this personal data for? What is the purpose of using the personal data?

- The selected biographical data is required to enable analysis of background information that corresponds with an increased level of awareness and use of accent and dialect under investigation. This will allow for a greater understanding of factors that contribute to maintaining community, regional and linguistic identity. Interview questions are designed to elicit language and attitudinal information to enable analysis and record examples of accent and dialect in use as well as awareness of this linguistic variety. This will increase understanding of the present use and awareness of this variety and contribute to maintainence of this variety and a general understanding of factors that promote or preclude maintainance of non-standard varieties of English and its contribution to regional linguistic varities.
- Attitudinal scales are designed to enable quantitative analysis of identity factors that correspond with an increase in awareness and use of the linguistic variety at the focus of the research project.

How are you making people aware of how their personal data is being used? Do you need to update your privacy notice?

- The consent form as well as project information sheet will present the details of how the data will be used and maintained.

Which conditions for processing apply for your project?

- For Special Categories please ensure you select at least one from Section 1 and one from Section 2 below.

Please select all that apply and provide any additional details.

Section 1: Conditions for Personal Data

The data subject has given consent to the processing (please provide the consent wording and where it is stored)

- data subject will give consent to process the data as outlined in the consent form. The wording is set out in the consent form.

Is all the personal data you are using necessary? Are you collecting enough to carry out the work, is there any?

you could do without to limit the risks to the individuals?

- The data is necessary to the purpose of the investigation and any subsequent data not required will be immediately erased.

How are you ensuring that personal data obtained from individuals or other organisations is accurate? How will you keep it updated?

- Data will be recorded with the participants present and will be clarified and checked where any issue of accuracy is a concern.

How long will you keep the data and how will you dispose of it? Are the retention periods on the University Retention Schedule?

- The data will obe archived for future research purposes and this will be stated in the consent form. The data will be anonymised once the research project is complete before submitting for archiving.

Where will the data be stored? If storage is in the cloud, where is the physical server? Will you need to transfer the data outside the EEA? If yes, how will you ensure adequate protection?

- The data will be stored on NTU's storage facility, and a password protected laptop used only by the researcher. On completion of the project, anonymised data will be archived and made available to affiliated researchers with legitimate and relevant research enquiries.

Will you be able to meet all the Data Subject Rights? Can you provide copies of data if requested? Are you able to fully delete the data (not just archive)?

- Data will be reproduced upon request for particiapants if requested. The data will also be deletable by the researcher if required, in order to meet the Data Subject Rights

Please briefly document below any risks with the use of personal data and how you will control such risks. Include technical controls (IT security, encryption etc), physical controls (location, locked room etc), personnel controls (training, access control etc), and procedural controls (contract, polices etc).

- The use of password protected laptop will be limited to the researcher.
- Transfer of files will be limited to researcher owned and protected equipment.

Appendix 8

Transcription Conventions

.	end of intonation unit: falling intonation
…	minimal pause
(1.5)	length of pause
[]	added word for clarification
1	line number
Bold	variant for discussion
Clinton:	Name of participant
Int.	interviewer
Sp2.	Other participant (see table 3.0 for interview context)

Appendix 9

Archive sources

21st Century

Acton, T. & Kenrick, D. 1984. Romani Rokkeripen To Divvus. *The English Romani Dialect and Its Contemporary Social, Educational and Linguistic Standing*. London: Romanstan Publications. **(AK).**

Dawson, R., 2002. *The Dialect of Derbyshire's Traditional Travellers*. Self-published. ISBN 1-903418-28-3. **(DV2).**

Dawson, R., 2009. *An English to Romany/Scottish/Irish Traveller Dictionary*. Self-published. ISBN 978-1-903418-66-6. **(DV5).**

Dawson, R., 2011. *An Irish Traveller Pavee Cant (Gammon/Shelta) Dictionary.* Self-published. ISBN 978-1-903418-78-9. **(DV4).**

20th Century

Dawson, R., Dawson Vocab 1959-69. Special Collection. University of Reading Special Collections and The Museum of English Rural Life. Collection reference D BD 1/8 – D BD 1/32. **(DV1).**

Dawson, R., Special Collection. 20th C. University of Reading Special Collections and The Museum of English Rural Life. Collection reference D BD 1/8 – D BD 1/32. **(DV3).**

East Anglian: No author. 1929. Anglo-Romani gleanings: from East Anglian Gypsies. *Journal of the Gypsy Lore Society*, Third Series, 8: 105,134. **(EA).**

Evans, I.H.N, 1929a. Gleanings from English Gypsies. *Journey of Gyspy Lore Society*, New series, 8 (3) 140-142. **(EV).**

Fox, Samuel. 1926. *The dialect of the Derbyshire Gypsies*. JGLS, 3rd series, 5, 62- 94. **(F1).**

Griffiths, J. & Yates, D. E. 1934. Sanderson's vocabulary, part 4: Rommano-English Vocabulary. In: *Journal of the Gypsy Lore Society*, third series, 13. 59-88. **(GY).**

Grosvenor, Lady Arthur. 1908. Whiter's Lingua Cinguariana. *Journal of the Gypsy Lore*. **(G).**

Hamp. No author. 1929. IV. *Anglo-Romani gleanings*: (III) Chiefly from Hampshire. JGLS, 3rd Series, 115-139. **(H).**

Leeds Special Collection. 1979. "Can you pooker Romany? Gipsy and Tinker cant of Britiain and Ireland." cassette tape. In: Gypsy, Traveller and Roma. Collections. BC MS Rom/Fraser/I/1/6. **(LSP).**

Lucas, J. 20th C. *Journal of the Gypsy Lore*. New Series. Vol 8 Part 4. in Robert Dawson Special Collection. **(L1).**

MacGowan, Alan. 1996. The Winchester confessions 1615-1616. Depositions of travellers, Gypsies, fraudsters and makers of counterfeit documents, including a vocabulary of the Romany language. South Chailey (East Sussex): Romany and Traveller Family History Society: in Manchester Corpus. **(MG).**

Manchester (Corpus) Romani Project. 2021. [online]. [Accessed 5 January 2018 - October 2021]. Available from: http://romani.humanities.manchester.ac.uk/ **(MC).**

Rieder, M,. 2018. Irish Traveller Language An Ethnographic and Folk-Linguistic Exploration first edition Palgrave Macmillan. **(EM).**

Roberts. No author. 1912. Robert's Vocabulary. *Journal of the Gypsy Lore Society, new series*, 5, 177-192: in Robert Dawson Special Collection. **(R2).**

Russell, A. 1915. Tom Taylor's Anglo-Romani vocabulary. *Journal of the Gypsy Lore Society*, new series, 8, 281,287. **(R).**

Sampson, J. 1911. Jacob Bryant; Being an analysis of his Anglo-Romani vocabulary, with a discussion of the place and date of collection and an attempt to show that Bryant, not Rudiger, was the earliest discoverer of the Indian origin of the Gypsies. *JGLS*. New Series, 4: 162-94. **(S3).**

Sampson, J. 1926a. *The Dialect of the Gypsies of Wales*. Oxford: The Clarendon Press. **(FS).**

Taylor, T. 1915. Word list. In Russell, A. 1915. Tom Taylor's Anglo-Romani vocabulary. *Journal of the Gypsy Lore Society*, new series, 8, 281,287. **(T1).**

Various Sources 1. 20th C. *Journal of the Gypsy Lore Society*, 3rd series, Volume XXVII Part 3-4 and XXVIII parts 1-2. **(V2).**

Various Sources 2. 20th C. *Journal of the Gypsy Lore*. 3rd Series. Vol II Part 4: in Robert Dawson Special Collection. **(V1).**

Way. 20th C. *Journal of the Gypsy Lore*. 3rd Series. Vol 34 Part 3-4: in Robert Dawson Special Collection. **(W1).**

Winstedt, E. O. 1948. Anglo- Romani gleanings from the northern counties. *JGLS*, third series, 27: 83–110. **(W2).**

19th Century

Borrow, G.A. 19th C. Romani Luvo Lil: in Robert Dawson Special Collection. **(B2).**

Bright, R. 19th C. *Journal of the Gypsy Lore. New Series*. Vol IX Part 3-4: in Robert Dawson Special Collection. **(B1).**

Crofton, B & Smart, H. 1875. *The dialect of the English Gypsies*. 2nd Edition. London: Asher. **(SC1).**

Irvine's Vocabulary. No Author. *Journal of the Gypsy Lore*. 19th C. 3rd Series 4. 172-182: in Robert Dawson Special Collection. **(IV).**

Norwood. No author. 19th C. *A vocabulary of words in Gypsie language or Romanish*. Norwood Notebooks: Scott MacFie collections: in Robert Dawson Special Collection. **(N1).**

Sampson, J. 1891. *Journal of the Gypsy Lore*. 1st Series. Vol 3 Issue 2: in Robert Dawson Special Collection. **(S2).**

Sanderson. 19th C. *Journal of the Gypsy Lore*. 3rd Series. Vol XIII Part 3: in Robert Dawson Special Collection. **(S1).**

Winstedt, E. O. 19th C. Anglo-Romani gleanings from the Northern Counties. *Journal of the Gypsy Lore Society*, 3rd series, 28. 50-61. **(W3).**