

Welsh Lexical Planning and the Use of Lexis in Institutional Settings

Elen Robert

A thesis submitted for the degree of PhD

Centre for Language and Communication Research
School of English, Communication and Philosophy

Cardiff University

December 2013

Summary

This thesis considers what I call lexical planning initiatives for Welsh – formal attempts to codify and standardise Welsh words. Welsh has been subjected to lexical planning – and purification – attempts for a number of centuries, with lexicographers seeking to coin and standardise Welsh-equivalent words for concepts that have initially emerged through contact with English (and other prestigious languages). Lexical planners have redoubled their efforts in the last fifty years, but especially since 1993, largely as a result of the language revitalisation movement. Lexical planning efforts can be envisaged as attempts to influence the acquisition and use of *any* lexical resources, but they often focus on specific subject matters, especially from modern or emergent domains or disciplines. Such initiatives are often referred to as terminology planning/standardisation.

My research considers the implementation of these planning initiatives, focusing on spoken language data at two research sites: the broadcast media and an office-based workplace. Taking a two-pronged approach to analysis, I ask whether, how and why Welsh speakers use planned lexis. First, I consider the extent to which the lexical content is in keeping with the stipulations of lexical planners in their codification texts. This approach is chiefly quantitative, drawing broadly on corpus linguistics and variationist sociolinguistics. Secondly, taking a more context- and practice-focused, as well as critical approach, I undertake an interaction analysis of the in situ use of lexical resources. From this perspective, we gain a picture of the underlying, sometimes conflicting, ideologies and discourse priorities that motivate lexical choice. This approach considers lexical planning initiatives not as implemented top-down, but embedded in their social milieu.

Finally, I consider the implications of my research for the broader revitalisation effort, asking to what extent lexical planning initiatives, as they are currently imagined and conducted, complement other language planning endeavours, and whether and how they might be reconsidered.

Statement 1

Except where indicated by specific reference, the work submitted is the result of my own investigation and the views expressed are my own.

Statement 2

No portion of the work presented has been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Statement 3

I hereby give my consent for my thesis, if accepted, to be made available for photocopying, inter-library loan, and in electronic format (PDF), and for the title and summary to be made available to outside organisations

Signed (Elen Robert)

Table of contents

<i>Figures and tables</i>	v
<i>Transcription conventions</i>	vi
<i>Acknowledgements</i>	vii
1. Introduction	1
1.1 Research context	1
1.2 Research approach	7
2. Language planning and standardisation	13
2.1 Introduction	13
2.2 Lexical planning, Terminology and lexicography	14
2.3 The aims and assumptions of lexical planners	19
<i>Standardisation versus codification</i>	19
<i>Technical registers versus standard language</i>	21
<i>Spoken versus written language</i>	24
<i>Effective communication versus social meaning</i>	26
2.4 Standardisation and standard languages	31
2.5 Implementation	30
3. Methods and data	47
3.1 Introduction	47
3.2 Ethical considerations	49
3.3 Media data: Dr Ann's <i>Syrjeri</i>	53
3.4 Workplace data: Housing Association	56
3.5 Subject-specific lexis as codified	59
<i>Selecting a sample</i>	60
<i>Sources consulted</i>	62
<i>Coding and collating</i>	66
<i>Critique</i>	72
3.6 In situ language policy/practice	75
<i>Social meaning and bilingualism</i>	79
<i>Social meaning and normalisation</i>	84
4. The Welsh sociolinguistic context	89
4.1 Introduction	89
4.2 A brief history of Welsh standardisation and lexical planning	89
4.3 Lexical variation and borrowing	98
4.4 Planners, products and procedures	108
4.5 The media context	116
4.6 The workplace context	124
5. Mediating medicine in Dr Ann's <i>Syrjeri</i>	131
5.1 Introduction: the <i>Syrjeri</i> data	131

<i>Structure and content of Syrjeri items</i>	131
<i>Discourse goals and priorities</i>	135
5.2 Subject vocabulary and codification texts	137
5.3 Codified lexis in <i>Syrjeri</i> talk	143
<i>Conclusions</i>	156
5.4 In situ language practice at the <i>Syrjeri</i>	159
<i>Conclusions</i>	190
6. Work talk at the Housing Association	193
6.1 Introduction: the Housing Association data	193
<i>Structure and content of Housing Association meetings</i>	193
<i>Discourse goals and priorities</i>	195
6.2 Subject vocabulary and codification texts	197
6.3 Codified lexis at the Housing Association	204
<i>Conclusions</i>	213
6.4 In situ language practice at the Housing Association	215
<i>Conclusions</i>	242
7. Conclusions and implications	245
7.1 Whether, how and why? Addressing my research questions	245
7.2 Methodological matters	250
7.3 So what? Implications for language planning and policy	254
References	267
Appendices	287

Figures and tables

Table 3.1: Example of information collated from <i>Geiriadur yr Academi</i> on the media dataset	70
Table 3.2: Example of information collated from <i>Termiadur</i> on the workplace dataset	70
Table 4.1: A small selection of words first attested in the dictionaries of Walters and Pughe	92
Table 4.2: Winford's framework of lexical borrowing	101
Table 5.1: Types and tokens of all terms in the <i>Syrjeri</i> sample	143
Table 5.2: <i>Syrjeri</i> term types according to codification category for participant groups	153
Table 6.1: Types and tokens of all terms in the Housing Association sample	205
Table 6.2: Housing Association term types according to recording and degree of codification	211

Transcription conventions

- (.) untimed pause
- (1.5) timed pause
- <text> following code applies to all text within triangular brackets
- [?] preceding is best guess
- (text) comments on context or delivery
- xxx unintelligible material
- text- truncated word
- text** item under discussion (or its equivalent in translation)
- text heavily accented syllable
- : elongated sound
- //...\\ enclosed talk overlaps with the line below
- /...\\ enclosed talk overlaps with the line above
- = quick take-up of following utterance
- #ty letter of the alphabet. Spellings seek to represent pronunciation (in this example, /tə/, the Welsh pronunciation of the letter 't')

These transcription conventions relate to extracts from the data as presented in the main body of the thesis. Conventions used in the sample transcripts in the Appendices differ marginally from these. I give more detail in the relevant Appendices.

I have typically sought to represent pronunciation where it differs markedly from spelling (e.g. /'kartra/ as *cartra* rather than *cartref*). I have transcribed prepositions as heard (e.g. *ona ni* ('of us') rather than *ohonon ni*).

I have sought to use Welsh orthography for borrowings, where I feel they are well-established and the sounds are easily represented. English orthography is used for other borrowings. I have represented /tʃ/ in Welsh orthography as 'ch' (e.g. *checio* for 'to check'), despite the letter usually being reserved for /χ/.

A translation is given under each line of transcription. I have included question marks in the translation where I interpret the speaker to be posing a question, as well as quotation marks where they seem to be quoting. I have also marked pauses and best guesses in the translation where possible.

Acknowledgements

Hoffwn gydnabod gyda diolch y nawdd gan Fwrdd yr Iaith Gymraeg a wnaeth yr astudiaeth hon yn bosibl. Diolch am bob cymorth gan staff y Bwrdd yn ystod y broses ymchwil.

Fasai'r ymchwil hwn heb fod yn bosibl heb gydweithrediad parod staff yng nghwmni Tinopolis, a hefyd yr holl gyfranogwyr, yn staff ac yn denantiaid, a gymrodd ran yn yr astudiaeth o'r Gymdeithas Dai. Diolch o waelod calon i'r unigolion ac i'r sefydliad am ganiatáu i mi gymryd cofnod mor uniongyrchol o'ch sgwrsiau ac am bob cymorth yn y broses.

Rwy'n ddiolchgar i staff a myfyrwyr ôl-raddedig Canolfan Ymchwil Iaith a Chyfathrebu (CLCR) Prifysgol Caerdydd am bob help llaw dros y pum mlynedd diwethaf. Mae diolch arbennig i'r Athro Nikolas Coupland ac i Dr Frances Rock, fy ngoruchwylwyr amyneddgar, am bob ysbrydoliaeth, cyngor ac ysgogiad – a hyd yn oed am droi tin min ar adegau priodol. Allwn i ddim fod wedi gofyn gwell.

Diolch i deulu a ffrindiau, yn arbennig fy rhieni am bob cefnogaeth ac am yr holl oriau o ofal plant yn y ddwy flynedd ddiwetha. Chi werth y byd. Yn olaf, diolch arbennig i Gruffydd am dy amynedd diflino drwy'r cyfnodau anodda, ac i Owain Siôn am lawenydd a phersbectif.

CHAPTER ONE:

Introduction

1.1 RESEARCH CONTEXT

In 1991 two Welsh-language campaigners were sentenced to 12 months in prison for causing damage to government offices as part of a campaign by *Cymdeithas yr Iaith Gymraeg* ('the Welsh Language Society'). Years later, one of the campaigners related to me how, on visiting a relative, he found him struggling to fill in the Welsh version of a form issued by the government, with its unfamiliar style and new-fangled words. Finally giving up on trying to decipher the form, the relative threw the papers at the campaigner, saying 'Here! This is what you went to prison for, isn't it?'¹ It is not unusual for government forms to be difficult to understand in any language. The *Plain English Campaign*² was set up specifically to call for greater clarity in all types of writing in English, but in particular in public-targeted writing. There is nothing particularly different about Welsh that somebody would find it difficult to fill in the Welsh version of a government form. The *Plain English Campaign* even has a Welsh-language equivalent in *Cymraeg Clir* ('clear Welsh') (C. Williams, 1999). But there is the added problem for Welsh speakers (and the authors of Welsh-language texts/translators) that there isn't a strong convention of using Welsh in certain contexts, including public administration. These contexts are called language 'domains' in the bilingualism and language planning literatures, (Fishman, 2007 [1965]). Consequently there is a perceived lack of well-established language practices, perhaps most obviously a lack of 'Welsh' words. Language planners see the problem in terms of lexical gaps in new domains. Prys (2006: 41), one of the most prominent Welsh lexical planners, summarises the problem thus:

¹ This episode was personally related to me by the campaigner.

² www.plainenglish.co.uk/

“[Minority] languages usually have very poor terminological resources. If they are not used for affairs of the state they will not easily develop the necessary vocabulary to deal with these matters, and may be excluded from domains such as public and private administration, science and technology, business and industry.”

Planners seek to fix the problem by coining words to fill these gaps and seeking to ‘standardise’ them (following their own definition of the word, see Chapter 2). This is referred to as lexical elaboration, modernisation or cultivation in the language planning literature, an activity that is central to corpus planning (Kloss, 1969)³ and language standardisation. Whilst planners take a problem-solving attitude to their work, the use of new words can add to the unfamiliarity of texts (written and spoken) and lead to further frustration, like in the example quoted.

Prys’ quote above also relates the problem of language style and competencies to efforts to revitalise the Welsh language, as well as to a language rights discourse (the line between revitalisation and language rights in relation to Welsh is typically blurred). Prys assumes that Welsh will continue to be excluded from certain domains unless words are coined and standardised to enable its use; status planning goals – the use of Welsh for more functions, especially higher-status functions (Fishman, 1991), such as public administration – are presented as the impetus for her corpus planning activities, a common status-corpus planning relationship (Fishman, 2006). The example of the campaigner, his relative and the government form also relates to the revitalisation/language rights effort, although it draws attention to problems raised by language planning, not problems solved. I should explain that the campaign in question

³ Kloss doesn’t go into much detail on corpus and status planning, but distinguishes between planning that is concerned with seeking to influence the structure or form of a language (its ‘corpus’), and planning that is concerned with a language’s status vis-à-vis another language. As Kloss acknowledges, the label *language planning* was originally used to refer to the former activity (as in Haugen’s (1972 [1961]) conception). It is also worth noting that *status* planning can refer either to planning a language’s social status (as in Kloss’ original definition) or to planning the functions that a language fulfils (Cooper, 1989). Fishman’s (1991) definition of status planning conflates both meanings.

called for a Property Act to tackle in-migration to Welsh ‘heartland’ areas and its effects on language demographics and local language practices. In itself, the campaign did not focus on status planning goals; the campaigner did not go to prison over Welsh versions of government forms. But status planning – especially targeting public administration – has been central to revitalisation efforts and to securing language rights. In fact, one of the most well-known Welsh-focused campaigns had a clear status planning goal. During their long-running campaign during the 1950s, which in some ways contributed to the founding of *Cymdeithas yr Iaith*, Eileen and Trefor Beasley refused to pay their rates to Llanelli Rural District Council until bills were issued in Welsh/bilingually⁴ (Phillips, 2000). This is the wider revitalisation/language rights effort that the campaigner’s relative, in the example above, was alluding to, part of the cause for which the campaigner purportedly went to prison. The consequences of status planning for revitalisation and language rights, as the relative understood, are somewhat grey. The use of Welsh in new contexts, which are assumed to necessitate a standard language, where new words have to be coined and old words refashioned, can effect a style of language that is difficult for speakers to penetrate. This is the realpolitik that the relative threw at the idealistic campaigner, metaphorically speaking; the result of language planning can be Welsh-language government forms that Welsh speakers don’t understand.

The Welsh Language Board⁵ rationalised its own work in what I will call lexical planning⁶ (lexical elaboration and standardisation) as providing “the linguistic resources needed to facilitate bilingualism”⁷ and “developing a form of Welsh

⁴ Often ‘in Welsh’ and ‘bilingually’ mean the same thing: as long as Welsh is present. Phillips writes that the Beasleys sought bilingual forms. But surely the political point they were making was for the presence of Welsh, and not specifically for bilingualism.

⁵ The Welsh Language Board was a government-funded public body that was set up to regulate the implementation of the Welsh Language Act 1993, and which was later given responsibility and resources for carrying out the Welsh Government’s language planning priorities. It was abolished in 2012. Its work is now carried out partly by the Welsh Government and partly by a newly established Language Commissioner. It is the Language Commissioner who has taken over the Board’s work in corpus planning.

⁶ I discuss my choice of the term ‘lexical planning’ in the next chapter.

⁷ Similar to note 4, the reference to ‘facilitating bilingualism’ here can be taken to mean ‘facilitating the use of Welsh’.

which is popular, useful and used” (Welsh Language Board, 2005: 36). Given what I have outlined above, we need to ask how popular and useful the language forged and codified by lexical planners in Wales is. Does planned standardisation, and lexical planning more specifically, facilitate the use of the Welsh language? And ultimately, does it help the revitalisation project? Is this securing language rights?

In relation to the language rights discourse it is a particular conundrum. Lewis bases his argument advocating the legally embodied right to speak Welsh on the value of language as the primary force in shaping identity: “Pwysleisir gwerth cynhenid iaith er mwyn tanlinellu pa mor ganolog yw iaith i hunaniaeth yr unigolyn” (Lewis, 2008: 8)⁸. Drawing on the ideas of Ignatieff (1994), he argues:

“Nid ydym yn llwyr gyfforddus mewn sefyllfa gymdeithasol [...] os na chawn ni siarad ein mamiaith: i brofi hapusrwydd llawn, rhaid i eraill ddeall nid yn unig yr hyn a ddywedwn, ond yr hyn a *olygwn*, a rhaid i ni fod yng nghwmni ‘ein pobl ein hunain’ i sicrhau hynny.”
(Lewis, 2008: 7, original emphasis)⁹

I don’t want to argue against the value of shared meaning-making, but the wider argument about language and identity here is based on the over-simplistic notion of the ‘mother-tongue’. In the face of sociolinguistic variation, what does ‘mother-tongue’ mean? What is ‘language’? This is a problem in all languages. Is an English speaker ‘completely comfortable’ in a courtroom, as defendant or witness, where they are being cross-examined by a solicitor who is trained to manipulate language for their own ends? Do they understand fully what is going on? Does the variety of language they use affect the way they are perceived by the jury? Does the rest of the courtroom understand fully ‘not only what they say,

⁸ ‘the intrinsic value of language is emphasised in order to underline the centrality of language to the individual’s identity’.

⁹ ‘We aren’t completely comfortable in a social situation [...] unless we are allowed to speak our mother-tongue: to experience complete happiness, others must understand not only what we say, but what we *mean*, and we must be in the company of ‘our own people’ to ensure that.’

but what they mean' just because they are speaking 'English'? There is a power differential at work that is not addressed by the 'right to speak my mother-tongue' argument. In the case of Welsh, there is the added problem of elaboration. If words have to be coined and standardised, and if speakers have to engage in situations where the language is unfamiliar to them, is it really plausible to assume that this particular version of language is central to their identity and is an aid in shared meaning-making?

It is a question that has arisen in other minority-language cultures¹⁰. For example, Wright (2007) presents the case of Arbresh speakers, who emigrated to Sicily from Albania 500 years ago. Recent status planning efforts, in particular to introduce Arbresh as the medium of education and as an educational subject, faltered in the face of the lack of a written standard, as well as disagreement when a standard was sought. The standard finally imposed is characterised by Wright as highly literary, difficult to understand and alien to the Arbresh speakers, since it is too far removed from spoken practices. This has caused a lack of commitment to the form, which has contributed to the slow implementation of the planning efforts, and called into question the value of these efforts in the first place. Similarly, efforts to standardise sign language in Germany and the UK in order to aid the teaching of sign have been met with suspicion by Deaf people (Eichmann, 2009: 301):

"If standardization was implemented because the existence of a standard language variety was valued in the majority society as a marker of 'real' language, Deaf people would be forced to change their language into something they do not recognize as authentic in order for the language to be acknowledged as genuine. This is seen

¹⁰ I use the term 'minority' guardedly, only to draw attention to the commonality between cultures, that "they are dominated politically and economically by numerically larger communities within a particular state" (Cormack, 2007: 2). Some researchers prefer to differentiate between minority (smaller in number) and minoritised language. The latter highlights minority status as the result of political forces which typically lead to the exclusion of the language from certain domains and institutions. See Drakeford & Morris (1998) for a more detailed discussion.

as inherently contradictory and highly contentious.”

It seems that if corpus planning is potentially contentious, it needs to be scrutinised, in particular since it is presented as problem-solving.

I don't want to assume that corpus planning necessarily causes more problems than it solves. It is the view of some that standardisation is a price that minority-language cultures have to pay if they are to be maintained, at least within a nation-state political framework (G. Williams, 1992: 147). Whether it is necessary or not, there is potentially a certain degree of inevitability that “discourses of resistance to centralizing states [such as minority-language activism] take up the legitimizing discourse of the state for their own purposes” (Heller, 2007: 5). In Fishman's view, standardisation, or corpus planning, doesn't come at a price at all (or at least he considers that the benefits far outweigh the negatives), as long as corpus planning is carried out well:

“The continuing opposition to corpus planning [...] can no longer successfully pretend that corpus planning cannot be done nor that it is impossible to do it well. It is, instead, ever more drawn into discussions of *who* should do it, of *when* it should be done, and of *how* it should be done, rather of whether it *can* or *should* be done at all.” (Fishman, 1983: 108)

I don't want to take Fishman's advocating stance on corpus planning, a stance he has made elsewhere (e.g. Fishman, 2006). I don't consider that the question of whether it 'should' be done has been answered, although I believe it is inevitable that it *will* be done within modern and post-modern societies, in particular in minority-language cultures. I try to take a value-free stance on corpus planning, although it is true that I come from a background where corpus planning was taken to be commonsensical and necessary, if not wholly unproblematic. Also, unlike Fishman, I try not to take a corpus planner's approach to the issue. I am

not concerned with ‘who should carry out corpus planning, when and how’. I try to take a sociolinguistic view of corpus planning as a cultural phenomenon and seek to investigate how it works, moves and influences (or not) in social space. Or, to reframe, I investigate how corpus planning is taken up within the social milieu in which it is embedded. This does not mean that the study offers nothing to language planners. Indeed, the study has been funded by the Welsh Language Board. But by approaching the issue from this perspective, I hope the study can offer a balanced view (as balanced as can be hoped for) of an activity that is often driven by personal interest and strong feelings (yes, corpus planning can be highly emotive).

1.2 RESEARCH APPROACH

We can approach the take-up of lexical planning in its social milieu from various perspectives. For example, my opening anecdote about the government form relates to whether planned language facilitates the use of Welsh, since in this particular context the propositional content of the message was not communicated well through the medium of Welsh, or at least through the language used in the form. The example contrasts planned language with the competencies of Welsh speakers (readers of the text in this case). Planning linguistic form is only half the story. If readers don’t understand the language, propositional meaning can’t be communicated. We might ask, then, whether planned language helps Welsh speakers communicate. This partly rests on whether they are (passively or actively) competent in the forms planned. On the other hand, if, according to Lewis’ (2008) ideas, above, concerning the need for securing language rights, planning is carried out primarily for identity reasons, we might ask whether Welsh speakers identify with the language planned. Do they relate to it? Do they feel they own it? Even the case of the government form is identity-related. Throwing the form can be seen as an act of rejection, not only of the form itself, and the language it was written in, but of the political struggle that underlay it.

These questions are also related to authenticity, an issue that was explicated in the Arbresh and Sign Language examples. Moreover, my brief discussion critiquing Lewis' argument points out that the uncritical linking of 'mother tongue', meaning-making and identity is problematic, especially since planned language doesn't fit easily within the concept 'mother tongue' since it is fabricated, often after the primary language socialisation stage. Moreover, unbalanced competencies in standardised language between interlocutors can (although do not necessarily) lead to one party gaining advantage and power over the other. We might, then, take a critical approach to the issue and investigate the role of planned language in the negotiation of power and authority. Is planned language used by speakers as a means of controlling events in order to achieve their aims, as Tollefson defines power (2006: 46)? Again, this question is related to the Sign Language example, since, according to Eichmann, standardisation is felt to be driven less by the needs and wants of Deaf people than by the Hearing majority.

As a starting point, we can consider the *use* of planned language by Welsh speakers. Language planning must have *use* as its main priority if it primarily involves "deliberate efforts to influence the behaviour of others" (Cooper, 1989: 45). Status, corpus and acquisition planning all ultimately target language use. Even approaches that target attitudes, ideologies or beliefs, such as planning for tolerability (de Bres, 2008; May, 2000) or prestige planning (Haarmann, 1990), ultimately seek to influence language use, if indirectly, since the goal is to enable or encourage certain language practices. Of course, corpus planning can be scrutinised by looking at the way corpus planners go about their work, the assumptions they make about the ideal language, about how the planned language should be disseminated, etc.. But without studying use, or language practices, we get no indication of its social impact.

The use of planned language has been the focus of some research on term

planning¹¹, which has typically referred to the use of planned terms as term *implantation*. For example, Quirion (2003) proposes what he calls a *terminometric survey*, which is “a study designed to measure the degree of implantation of all designations referring to a single concept or to a set of concepts” (2003: 30). His terminometric survey focuses exclusively on the extent of use of planned terms within (written) institutional communication (that is, *whether* planned terms are used), thereby offering a quantitative account of the ‘success’ of term planning efforts within a given polity. Whilst investigating *whether* planned language is used or not is undoubtedly central to researching the take-up of lexical planning, such an account wouldn’t tell us much about the interactional and social implications of corpus planning, which is one of my central concerns. For example, it won’t tell us whether the planned language facilitates the use of the language as a whole, nor about whether speakers feel ownership over the language, or feel it to be authentic, or what role it might use in the negotiation of social relations. As Fishman (1983: 108) argues, in relation to lexicons in particular:

“they are not endless laundry lists, without rhyme or reason, without systematic links to each other and to all other facets of language. Lexicons are not interchangeable, dry and dreary ‘nuts and bolts’ [...] their successful planning involves tremendously complicated socio-cultural-political sensitivities that most linguists neither possess nor imagine.”

Only considering *whether* planned language is used would be to ignore the socio-cultural-political embeddedness of language planning. Moreover, the term *implantation* itself, drawing on a metaphor that is almost surgical, seems to consolidate this decontextualised view, representing the phenomenon as a top-down procedure where those affected play no active role in the embedding of

¹¹ See the next chapter on the slight difference in focus between what I call ‘term planning’ research and ‘language planning’.

lexical planning. The term sits uncomfortably within a sociolinguistic framework, especially one that assumes that speakers have some level of reflexivity in language practice.

Within a sociolinguistic framework, we must take a more nuanced approach to language practice, and take sociolinguistic variation as a given. Planned terms are likely to be differently used across different types of interaction and for different purposes, as Ní Ghearáin (2009) found in the case of Irish: while all of her 37 participants expressed support for lexical planning efforts for Irish, seeing it as necessary for revitalisation, their more implicitly expressed attitudes were not as strong, and their own reported practices were variable. For example, they reported that they used English borrowings when they didn't know the Irish word, and didn't usually find this problematic, at least not in terms of the communication of propositional meaning, but usually not in terms of the performance of identity either. It was in the context of their worklives that they felt the need for planned terms most strongly, where they feared lack of competence in planned terms might reflect badly on their professional identities. However, even at work, it was only for some kinds of social activity that competence in planned Irish terms was considered important, specifically in writing and in formal spoken communication, such as presentations and speaking on the radio.

Instead of considering *implantation*, focusing solely on the uptake of certain words without consideration of how or why they are used or not, I will ask about the *implementation* of lexical planning. I go into more detail on implementation in Chapter 2, but briefly, implementation is a term taken from the language planning literature which opens the door to considering the nuances of language practices, such as those Ní Ghearáin picked up on. Coupland & Kristiansen (2011: 21) define implementation as “‘spreading’ processes, which have both an ideological aspect (acceptance) and a use aspect (diffusion)”. ‘Spreading’ processes relate to top-down measures of implementation as well as on-the-

ground use:

“it is likely that the process of standardisation will be understood quite differently by those engaged in top-down agentive roles and by others, ‘the people’, who make on-the-ground assessments of the social implications of using different ways of speaking. Top-down discourses of language standardisation may not overlap with on-the-ground discourses, and the social judgements of language use that matter most may even remain below the level of metalinguistic formulation” (2011: 22)

Whilst my research does seek to consider whether planned language is used, I also consider the discursive or ideological component of language use, considering whether top-down discourses do overlap with on-the-ground discourses, and if they do, when.

Finally, my research does not focus exclusively on the take-up of planned language. Spolsky (2004: 26) points out that, typically, language policies have two sides: one side seeks to forge new linguistic forms, according to certain criteria of what is good or desirable, and the other seeks to restrict the use of language that is considered unfavourable, inappropriate or bad. At the same time as seeking to forge a new language, lexical planning seeks to restrict another way of speaking. Indeed, standardisation is fundamentally a centralising discourse. I will argue in the following chapters that underlying Welsh lexical planning is an ideology of purism that seeks to restrict or control the use of borrowings in Welsh-language interaction (compare with Ní Ghearáin’s participants, who used an English borrowing when they didn’t know the Irish word). In considering the take-up of lexical planning, then, my research inevitably deals with bilingualism and language contact, and the ideology of purism.

My study is an investigation of naturally-occurring talk in two institutions, one a

media institution and one a workplace¹². The following two chapters lay out the conceptual and theoretical groundwork for my study of the take-up and social embeddedness of lexical planning. In Chapter 2, I investigate the key concepts under consideration and seek to pin down my own key terminology, including standardisation, lexical planning, terminology, codification, and implementation. I do so primarily within a discussion of what it is lexical planners, in Wales in particular, are seeking to achieve. In Chapter 3, I set out my methodology for the study. This includes: explaining my choices of data site, in terms of type and specific site; discussing the ethical considerations of data collection, analysis and presentation; giving some details of the data collected at each site; and detailing the approaches I took to analysis. I approach analysis from two different perspectives. First, I seek to offer a quantitative account of whether participants use planned lexis within the interactional episodes collected. This includes breaking the data down according to extralinguistic variables. Second, I offer a qualitative account of the data, through an interaction analysis, considering whether the interaction overlaps with the ideology and priorities of lexical planners. The analysis is dependent to some extent on sociolinguistic background knowledge, which I offer in Chapter 4. This includes more detail on lexical planning and standardisation in Wales and a sociolinguistic background to the institutions studied. Chapter 5 and 6 present my analysis of data taken from each site, each chapter investigating the take-up and embeddedness of lexical planning within one institution. In Chapter 7, I draw both analyses together, and discuss what implications there may be for the lexical planning project in Wales.

¹² Talk produced for television broadcast is naturally-occurring in the sense that it is not produced specifically for the research exercise.

CHAPTER TWO:

Language planning and standardisation

2.1 INTRODUCTION

The phenomenon under investigation in this thesis can be viewed through various lenses. Whilst the context of the work is language policy and planning (LPP), there are other research traditions that the research might contribute to, and that I have drawn extensively on. This chapter discusses the various research traditions and how they relate to the research. I start by considering the nature of the phenomenon under investigation, which is differently understood from various research perspectives. I have referred variously in the preceding chapter to the phenomenon using a number of designations: corpus planning, lexical planning, elaboration, modernisation, cultivation, standardisation and terminology. To some extent, these designations reflect a somewhat different understanding of the nature of the activity of planning. Lexical planners in Wales often refer to their work as *safoni termau* ('standardising terms') or *terminoleg* ('terminology'). Similarly, others outside of Wales researching efforts to forge and standardise modern lexicons, in particular for minority languages, and with a view to supporting these efforts, make some reference to *terminology*: either *terminology/terminological planning*, *term planning*, or *terminology policy* (Bhreathnach, 2009; Colomer & Cuadrado, 2007; Drame, 2007; Mwansoko, 1993; Ní Ghearáin, 2008; Quirion, 2003). My work largely addresses and draws from the Terminology¹³ and term planning literature(s), especially since so much corpus planning in Wales happens under the rubric of 'terminology'. However, I prefer to talk more broadly about *lexical planning*, since lexis-focused corpus planning in Wales has, in a number of ways, a wider scope than what is implied by 'terminology', which I outline below. As regards the other designations, corpus planning, elaboration, modernisation and cultivation are all terms used within

¹³ I capitalise 'Terminology' in order to differentiate the discipline from terminology as a perceived body of concepts and designations (the focus of the discipline of Terminology). See section 2.2.

the language planning literature. Corpus planning is superordinate to lexical planning; elaboration, modernisation and cultivation are generally subsumed by it¹⁴. Standardisation is a designation that deserves its own discussion, since its meaning is difficult to pin down, and since it has been such a central concept in sociolinguistics. I consider standardisation, below, as a label used by term planners and terminology researchers. However, lexical planning can also be considered through the lens of the research literature on standardisation. This literature gives us a slightly different perspective. I consider standardisation in more theoretical detail, then, in section 2.4, after discussing some of the priorities of lexical planners and the means by which their work is disseminated and implemented.

2.2 LEXICAL PLANNING, TERMINOLOGY AND LEXICOGRAPHY

With reference to ‘terminology’, I should explain here that there is a discipline of Terminology, which overlaps with the language planning literature but which has developed separately, and which largely sees itself as distinct from the language planning literature and from linguistics. It is relevant to this discussion because it provides a methodological approach to a lot of the lexis-focused corpus planning activities in minority language contexts. The discipline of Terminology originated in a science and technology context, and concerns itself with the study of the vocabulary of specialised fields of discourse, or Language for Specific/Special Purposes (LSP). The vocabulary of a particular subject-area is treated as determinate, leading to a simple view of the vocabulary as “the entirety of all concepts and terms in one specialist field”¹⁵. Eugene Wüster, an Austrian engineer, developed his original theory of terminology in the 1960s based on his experience of working on a French/English dictionary of technical terms (Cabr  Castellv , 2003). His aim was to improve scientific communication, and thereby further the progress of modernity, through imposing uniformity on the

¹⁴ Elaboration is also used to refer to planning elaboration of function/domain use (Kaplan & Baldauf, 1997: 43), but I use it here mainly to refer to the process of forming and disseminating lexical resources presumed necessary for the spread of a language to new domains.

¹⁵ *German Terminology Portal*, www.iim.fh-koeln.de/dtpEN/terminologieDTEN.html.

specialised lexical forms used in scientific discourse. Wüster brought with him to the field standardisation as an ideal which, in the social climate of the day in Western Europe, was driven by the globalisation of science and technology and ultimately, modernisation.

Whilst originally Terminology did not have a minority language focus, today much of the work on lexis-focused corpus planning in minority language contexts orients to this discipline. Methodologically, the practices of a number of minority-language lexical planners, at least in Wales, seem to draw more on Terminology than on the language planning literature or lexicography. I call planners taking this perspective *term planners*. For example, in their guidelines on how to carry out terminology projects Prys & Jones (2007) draw extensively for methodology on the standards developed by the *International Organization for Standardization* (ISO), or rather, by its dedicated Technical Committee, the ISO/TC 37 ‘Terminology and other language and content resources’. Other minority language lexical planners take a similar approach. *Minority Languages and Terminology Policies* was the subject of a conference held in 2007, which was organised by the *European Association for Terminology*. In their contribution on Catalan terminological planning, Colomer & Cuadrado (2007: 103) rationalise that “the ultimate goal of terminological work, in the fields of both translation and standardisation or documentation, is to foster communication between specialists and an exchange of scientific and technical knowledge in society”.

There is, then, an overt focus in the work of term planners on language for technical registers that serves communicative effectiveness (narrowly defined in terms of propositional meaning), following the broader Terminology literature. However, Drame (2007, 2008) chooses to distinguish between terminology in specialised fields or professional environments and terminology as part of corpus elaboration and modernisation in (national) language planning. Although there is considerable overlap between both types, it is a useful distinction to make, and I will differentiate between them lexically by reserving ‘terminology

management’ for the former (in specialised fields) and ‘terminology planning’ or ‘term planning’ for the latter (for national language planning).

In the context of Welsh lexis-focused corpus planning, however, I prefer to talk more broadly about *lexical planning*. I will sometimes refer specifically to terminologists, or term planners, because of their quite specific methodological orientation to language planning and since, through their work, they form a network of practice that can be considered, in some ways, different from that of lexicographers. But the phenomenon under investigation in my research is broader than this and is embedded in a culture of standardisation. As regards lexicography, I should explain here that the discipline of Terminology typically views terminology as distinct from lexicography (Cabr  Castellv , 2003). One difference is that Terminology allows fewer variants than lexicography, catering as it does for registers. This is apparent when we compare the products of lexicography and Terminology, namely dictionaries and glossaries¹⁶. For example, the English-Welsh dictionary, *Geiriadur yr Academi* (‘the Academy Dictionary’) (Griffiths & Jones, 1995), a lexicographical work, lists four Welsh forms under the English entry *jaundice*, some of which are marked as regional, occasional or archaic (they are *clefyd melyn*, *clwyf melyn*, *cric melyn* (‘north-west, occasional’) and *rhifwnt [melyn]* (‘archaic’)). On the other hand the English/Welsh glossary, *Y Termiadur* (Prys, Jones, Davies, & Prys, 2006), a terminological work, gives only one Welsh form (*clefyd melyn*, literally ‘yellow disease’). The difference arises from the latter’s goal of stipulating one and only one designation for each concept, following W ster’s standardisation ideal.

To return to my main point, I prefer to take a broader view and refer to *lexical planning* as a general framework for referring to and understanding the phenomenon under investigation. There are a number of reasons for this. Firstly, lexical planning in Wales is largely (if not exclusively) focused on lexical

¹⁶ By ‘glossary’, I mean a list or dictionary of terms, including online searchable databases. This is the word I mostly use for the products of term planners.

elaboration – the coining of new words for new concepts or new domains of use. Whilst lexical elaboration in Wales is largely carried out by term planners, or ‘terminologists’, I argue that elaboration is not uniquely within the scope of term planners, but also of lexicographers. Secondly, as we shall see in Chapter 4, lexis-focused corpus planning has a long history that extends back to the 16th Century, with elaboration beginning in earnest in the 18th Century (Morgan, 2002). Current ‘terminology planning’ in Wales clearly follows in this tradition. It takes a more subject-specific approach, coining and ‘standardising’ since the mid 20th Century lexis in subject areas such as history (University of Wales Press (UWP), 1959; Welsh Joint Education Committee (WJEC), 1987), economics (UWP, 1972; WJEC, 1992b) and information technology (MEU Cymru, 1992; Prys, Davies, & Prys, 2005; WJEC, 1986), to name but a few. But the main difference is not goals (the language planned and the uses planned for) but methods – how planning projects are carried out.

A final reason has to do with the overtly stated aim by Terminology researchers that term planning assumes a focus on scientific language within focused fields of discourse. In pursuit of this aim, ideally terminologies should be internally cohesive, with every label designating a single specific concept (what terminologists call *monosemy*, that is, the absence of polysemy) and every label designating only one concept (what they call *univocity*, that is, the absence of synonymy) (Cabr  Castellv , 2003)¹⁷, a method lexical planners in Wales subscribe to (Prys, 2007b). However, there has been lexis-focused planning in Wales ranging from highly technical IT terms (e.g. Prys et al., 2005) through to canoeing (Iorwerth, James, & Williams, 2007) and shop signs (Prys, Davies, & Prys, 2006). Can all of these reasonably be described as enabling ‘the exchange of scientific and technical knowledge’? And how important is it, realistically, to ensure the internal cohesion of these vocabularies? The dictionary of shop signs is clearly targeted at the general public. The point is to enable and normalise the

¹⁷ There have been critiques of these ideals from within the Terminology literature (e.g. Temmerman, 1997).

presentation of ('correct') Welsh in highly visual spaces – *everybody* shops. Canoeing is, of course, focused on leisure, which hardly fits the description given to terminology. Even in the case of the IT terms, I wonder whether they are intended only for experts or also for laypeople. The significance of information technology is that computers and other computer devices (smartphones, tablets, etc.) have radically changed the way we communicate. The translation of IT terms enables creating Welsh-language interfaces, and thus normalising the use of Welsh – for ordinary users (that is, *everybody*) as well as for experts. All of the dictionaries of terms I have mentioned seem to be focused on extending the use of Welsh to domains of direct social reproduction (Williams, 1992: 147) – planning for the general, not for the specific. I don't deny that there is also planning that targets expert communication and that seeks to standardise terms in highly specialised fields, where the internal cohesion of terminology sets is essential. This includes, for example, legislative terms, such as the *Highways Legislation Terminology* (Prys, 1998). But lexis-focused corpus planning in Wales is much broader than this.

It is worth noting here that there is a field of research that calls itself *socioterminology*, a field that has mostly been conducted to date through the medium of French. Most notable is the work of François Gaudin (2003), who expressly seeks to encourage a social perspective on terminology and specialist language. This involves, for example, understanding the terminology of a specific field in its socio-historical context, as part of the social construction of knowledge; a re-examination of the nature of the term, based on sociolinguistic understandings of domain; and considering terminology, primarily through interactional analysis, from the perspective of how speakers negotiate meaning – that is, taking a dialogic view of language (2003: 17-19).

Nonetheless, in relation to the scope of my study, much of the reasoning above holds. That is, I take 'lexical planning' to be a more appropriate designation than 'terminology' despite Gaudin's (and others') use of 'terminology' from a

sociolinguistic perspective.

2.3 THE AIMS AND ASSUMPTIONS OF LEXICAL PLANNERS

There are a number of priorities, assumptions or goals that lexical planners might aim for, which are pertinent to the question of the take-up of lexical planning. Some of these are overtly stated by lexical planners, others are covert and can be inferred from texts about lexical planning and from lexical planning products. These stances need to be clarified in order to understand the nature of the lexical planning, although often we will find that different planners have different objectives. Moreover, how planning efforts are taken up may differ from what planners originally intended. Nonetheless, in this section I consider the overtly-stated goals of standardisation, planning for technical registers, planning for written language, and planning for effective communication.

Standardisation versus codification

As I noted above, lexical planners in Wales sometimes refer to their work as *safoni termau* ('standardising terms') (e.g. Prys, 2007a; Prys & Jones, 2007; Prys & Jones, 1998a). But what is meant by 'standardisation'? It is a rather elusive term, itself testimony to the difficulties of insisting on monosemy and on the determinacy of fields of knowledge. It seems to be differently understood by Terminology researchers and by sociolinguists. When these literatures overlap (which they do in LPP research), there is some confusion.

From texts written by lexical planners, term planners in particular, it seems that standardisation is primarily conceived of as an activity that is carried out by language planners when they are compiling dictionaries and glossaries. This is seldom stated explicitly, but can be inferred, for example, when Prys (2006: 48) lists the "basic criteria for standardizing terminology" as "linguistic correctness; reflecting the characteristics of the concept; conciseness; ability to produce other forms; and monosemy". These criteria relate standardisation to a closed list of terms, and not to language practices. Standardisation in this sense is a statement

of the relationship between a concept and a carefully-selected label, a conceptualisation similar to that used by term planning researchers. For example, Bhreathnach (2009, 2012), drawing on the work of a number of Terminology researchers and international terminology organisations, lists a number of stages in term planning, including standardisation, which she places between research and dissemination. She defines standardisation as “the selection by a representative standardisation committee of recommended terms to be used in a defined field, such as in education or administration” (2012: 100). Standardisation is something lexical planners do.

Standardisation in this sense is closely bound with codification – “the explicit statement of a norm” (Cooper, 1989: 144). In fact, in his model of language planning, Haugen glosses codification (one of his four stages of language planning) as ‘standardization procedures’ (Haugen, 1983: 275). In codification, corpus planners decree the meaning of forms and the forms that should be used to express meanings. This involves the explicit suppression of variation by ruling out other designations, other realisations of the same basic form (e.g. different spellings) or other semantic meanings¹⁸. This perspective chimes with J. Milroy’s definition of standardisation – the suppression of variation (1999: 26), or in other words, “the imposition of uniformity upon a class of objects” (Milroy, 2001: 531). Arguably, term planners have a stronger drive towards standardisation than do lexicographers; term planners seek to define and impose a relationship between a single concept and a single label, whereas bilingual lexicographers often list a number of target-language variants under a source-language entry, as I explicated above with *jaundice* (although see my discussion on register, below). On the other hand, the practice of marking some of these variants as regional, occasional, familiar or archaic also works towards suppressing variation, since they are singled out as non-standard/-legitimate, and those that are not explicitly marked are positioned as ‘normal’ (Bourdieu, 1991: 90).

¹⁸ The suppression of variation in semantic meaning (that is, suppression of polysemy) is more prevalent in Terminology than lexicography.

To return to my main point, it seems clear to me that standardisation is more than something lexical planners do. If lexical planners are the agents of standardisation, then their products – dictionaries and glossaries (which I refer to collectively as *codification texts*)¹⁹ – are collections of standardised words. Indeed, lexical planners may well see it this way. But this can only be true in a narrow sense. Anybody can codify a list of words – make a statement as to the relationship between forms and meanings. But in what sense does this entail standardisation? Arguably, without the wider take-up of the linguistic forms codified, there is no standardisation. If standardisation entails the imposition of uniformity, that doesn't mean much if the language made uniform is confined within the boundaries of the codification text. What is the scope of this uniformity? I would argue that standardisation ought to be envisaged as a process of change – of decreasing variation – at the level of practice. The language selected and codified has to 'get out there'. By my understanding, standardisation is something lexical planners *aim for*, not an inherent quality in their work. Nor are the lists of words they codify standardised. This is why I call them codification texts, and not dictionaries/lists of standardised terms. I will consider standardisation in more detail below, in particular in relation to the LPP and sociolinguistics literatures.

Technical registers versus standard language

Related to standardisation and codification is the matter of what kind of standardisation lexical planners have in mind, according to the scope they envisage for their work. We can hypothetically differentiate between planning for standardised registers and planning for a standardised unified language, in its broadest sense. A maximally standardised register (e.g. a legal register) would be one where all those dealing with a particular field of interest on a regular basis use the same conventionalised language that is specific to that field. This kind of

¹⁹ Codification texts include tangible texts such as paper dictionaries, as well as online dictionaries and glossaries, and the databases that increasingly underlie them.

register is similar to the conventionalised linguistic resources that are assumed to develop within communities of practice²⁰. A community of practice is defined as “An aggregate of people who come together around mutual engagement in an endeavour” (Eckert & McConnell-Ginet, 1992: 464). In the process of engaging in a shared endeavour, over time the community of practice develops a shared repertoire of resources for making meaning (Holmes & Meyerhoff, 1999). These shared resources can be linguistic, which includes specialised terminology (arguably, a community of practice might share a number of different registers). An ideally standardised *language*, on the other hand, is theorised by Haugen (1966: 931) as one where there is minimal variation in form (which he calls codification) and maximal variation in function (which he calls elaboration). That is, it is one where all speakers of the language speak (and write) in the same way all the time for all functions. Absolute language standardisation, by this definition, is highly unlikely to be achieved, given the preponderance of sociolinguistic variation.

I consider the difference between these hypothetical types of planning because it potentially impacts how we should judge lexical planning (and in particular overtly register-oriented term planning). If planners target technical registers, does this come within the scope of ‘standard language’, and is it fair or necessary to consider how planned language is taken up by those who are not part of the community of practice? And yet, if we make the distinction between technical register and other types of language use, how do we judge whether an instance of talk is ‘technical’ or not?

The overt goals of term planners, as I was told by a terminologist in personal communication, is to target the users of technical registers, the implication being that planning for a unified, standard language does not come within the scope of term planning. In their guidelines on standardising Welsh terminology, Prys &

²⁰ Agha (2000) points out that there are other kinds of register, such as registers of respect or etiquette. Here I deal only with registers associated with fields of knowledge or with an organised activity, both of which can be considered within the scope of the community of practice.

Jones (2007: 43) make a similar statement:

“Terminology handling and standardization are aimed at the need for effective communication and knowledge exchange in technical registers of different subject fields.”

However, there is some reason to be sceptical about this overt goal, primarily because language planners orient to ideas about ‘correct’ Welsh. It may be that planners do not envisage the use of their planned language outside of certain fields or communities of practice, but as Schiffman argues (1992: 5), presumably with *technical* registers in mind, “Registers must be an integral yet specialized part of a recognized standard language”; technical registers are built from ‘standard’ resources. They come within the scope of standard language in a way that other types of register, such as Cockney rhyming slang, do not. This point is made clear in Prys & Jones’ guidelines on standardising terminology, where they stipulate that ‘standardised’ terms should be linguistically correct (Prys & Jones, 2007: 37). In its application, this is defined as conforming to certain codification texts (e.g. codified grammars) and giving precedence to native resources over borrowings. That is, planning for a technical register also involves, to some extent, planning for a standard language.

Moreover, as I have already discussed, the nature of lexical codification texts, subject-specific glossaries in particular, suggests a scope that is wider than technical registers or the register of a community of practice, if not by intent, then necessarily in its application. I noted that lexical planning projects have covered fields such as canoeing, IT and shop signs. What is technical about shop signs? These are terms for the ‘general public’, not subject specialists. Even in the case of IT, as I have argued, there is a clear concern with planning language that has a wider scope than subject specialists’ usage or the usage of a close community of practice. And even if we do accept that the goal of lexical planners, term planners in particular, is to plan for technical registers, since corpus

planning in Wales is so clearly bound with status planning (as I argued in the previous chapter), and since status planning in Wales places such emphasis on the provision of *public* services in Welsh, there is considerable crossover between technical registers and more general usage. So even if lexical planners aim to standardise technical registers, the language planned will inevitably surface in less specialised use, and this is built into the language planning framework that is the *raison d'être* of lexical planning.

Spoken versus written language

Within the Terminology literature, what is called 'standardisation' is assumed to target written language, not spoken, even when the goal is minority language planning (Cabr  Castellv , 2003: 168). Similarly, in the language planning literature it has been assumed that corpus planning primarily targets written language (Cooper, 1989; Fishman, 2006). However, corpus planning can also target spoken language. Kaplan & Baldauf (1997) have critiqued Haugen's depiction of corpus planning as *graphisation*, *grammatication* and *lexication*, since it is writing-centric. They cite Thomas' article on the development of a spoken standard for Welsh (Thomas, 1987) as evidence to the contrary.

It isn't clear how lexical planners themselves view their work. It seems that if the priority of term planners is ostensibly technical registers or 'special language domains' (Prys & Jones, 2007: 7), this isn't restricted to written language. For example, sport is a sphere where there is some dependence on a fixed relationship between lexical forms and meanings, but where communication is largely spoken (in as far as communication happens through language). In tennis, for instance, the words *game* and *match*, despite being very similar in denotational meaning, are used to designate different concepts. As technical, tennis-related terms, they are used in speech just as much as in writing. In my own analysis (in both analysis chapters), we shall see that there is a crossover between spoken and written communication, with participants sometimes, or even often, drawing on written language in the production of speech. It is not

clear where writing ends and speaking begins, and so ‘technical register’ does not easily map onto one or the other.

Moreover, Prys & Jones (2007) envisage the ‘vehicles of dissemination’ of their planned terms as including spoken media of communication, such as school lessons, radio and television. On the other hand, it is possible that there is an assumption that lexical planning isn’t particularly relevant to speech. In a meeting with a researcher on language awareness in healthcare provision, who had worked on a lexical codification text in the field, I asked about the role of lexical planning in healthcare. The researcher answered that Welsh isn’t used much for writing in healthcare, the unstated assumption being that lexical planning was only relevant to written language.

Whatever the answer regarding the overt goals of lexical planners, any corpus planning activity is likely to influence the linguistic culture more broadly, in particular in metalinguistic and ideological terms. Liddicoat (2000) claims that the 1966 publication of the *Dictionnaire Jersiais-Français* effected an ideological change in speakers of Jersey Norman French. Although the dictionary was intended to be a tool in the creation of a culture of Jersiais literacy, in effect, Liddicoat claims, it served as an authoritative statement of the language in its entirety, written and spoken, and of the competencies that the speakers of the language should (but in no way could) have. The consequence was a widespread belief amongst speakers of Jersiais that they did not know the language properly. The point as regards my own research is that whether or not planners intend for their work to be implemented in speech, codification texts may be seen as repositories of ‘the language’, and possibly affect the way speakers think about language, both written and spoken. Consequently, they have the potential to affect spoken practices.

Above and beyond these points, however, there is a danger of overconsolidating the difference between spoken and written practices. There are doubtless

differences resulting from the differentiated ways basic competence in speech and writing is acquired, and from the differentiated ways speech and writing must be articulated. However, the revolution in communications technology in the last 20 years suggests that beyond these points, what has previously been perceived as the unalterable differentiated norms of spoken and written media themselves is probably more aptly ascribed to the various social activities that are carried out through language and the conditions that enable and constrain them. On text messaging, for example, Thurlow & Poff (2013) argue that practices emerge from a combination of technological affordances (what practices are enabled by the technology) contextual variables and interactional priorities. Surely this is true of any communicative medium, even if it is particularly pertinent in the case of social media practices. E-mail, text messaging, blogs, social networking sites, etc. have allowed writing to be carried out under conditions vastly different from previously, which has undoubtedly influenced written form in many cases. This suggests that, with the exception of the difference between enunciation and orthography, variation between spoken and written language norms is largely an effect of differences in the conditions of the social activity that is lived through language practices. In terms of lexical planning, despite an overt focus on written language, there are likely both writing and speech events where planners envisage the use of their planned language and where speakers, as social actors, might orient to the linguistic and communicative norms endorsed by planners. As regards my own research, despite some claims that lexical planning targets written language, it is legitimate and worthwhile to consider the take-up of lexical planning in spoken practice, as I will do in my analysis.

Effective communication versus social meaning

As I have already noted, terminology planners tend to overtly consider their goal to be effective communication amongst experts, construing their efforts as attending to the propositional function of language. To give an idea of what planning with an eye on ‘effective communication’ might entail, Drame (2008)

gives the example of an aircraft maintenance business, which, she stresses, needs to operate in a highly multilingual environment where smooth and efficient communication is essential. She envisages that inconsistent terminology might lead to workers not being able to locate a particular mechanical part, for example, due to variability in its lexical designation. This might ultimately lead to the aircraft being held up and consequently to a disruption to the entire complicated system of flight schedules.

Although there may be some validity to this argument, the effective communication argument seems generally overstated, even in the case of terminology management (as opposed to planning). The following example highlights that terminology management can be driven by performative concerns that relate to the social meaning of using variable lexical resources. IBM's website has a section marketing its terminology management systems for aspiring global businesses, which stresses the need for language to be carefully controlled in company-customer interaction. The reasoning is, partly, that "Consistent terminology contributes to presenting an integrated look and feel across products"²¹. That is, (in)consistent terminology can have implications for the construction of corporate identity – whether the company is seen as an entity, and how that entity, as a corporate social actor, is perceived by others²². This suggests that Drame's characterisation of term management as attending to propositional meaning is somewhat oversimplified²³. Contemporary organisations typically seek to present single identities in their external communications, a task made more difficult by the multiplicity of social actors that produce the texts (spoken and written) that address recipients external to the organisation (since organisations are collectives). A multiplicity of social actors by necessity entails a multiplicity of voices, in a Bakhtinian sense. In

²¹ IBM, www-01.ibm.com/software/globalization/topics/terminology/introduction.html

²² Whether an organisation has an identity before the law is determined by its constitution, but organisations need to construct identities in relation to their (potential) clients.

²³ Other parts of the text emphasise that inconsistent terminology can cause customers to become confused regarding how to use products, and can even cause products to malfunction. These are hypothetical examples of the ineffective communication of propositional meaning.

Bakhtin's view, all words have histories, and, very simply put, give off those histories (or potentially do so) when used:

"All words have the "taste" of a profession, a genre, a tendency, a party, a particular work, a particular person, a generation, an age group, the day and hour. Each word tastes of the context and contexts in which it has lived its socially charged life; all words and forms are populated by intentions. Contextual overtones (generic, tendentious, individualistic) are inevitable in the word." (Bakhtin, 1935 [1981]: 293)

From this perspective we can see how the use of 'inconsistent terminology' might give off variable 'contextual overtones', and is potentially suggestive of multiple authorship and multiple intentions.

As regards lexical planning, there is *some* acknowledgement amongst planners that social meaning plays a role in codification decisions, although the concept of social meaning itself is not typically invoked. For example, in the section of their guidelines for terminology projects where they discuss whether borrowings should be used to create terms, Prys & Jones (2007: 37) write that:

"[In Welsh] there is often prejudice against borrowing from English [...] The aim of the international standards is to promote effective communication, rather than to preserve linguistic purity, but giving due regard to the wishes of the specific linguistic community is part of any comprehensive terminological planning."

Notwithstanding the problem of how lexical planners presume to speak on behalf of the 'linguistic community', here we see that social meaning is relegated to second place, behind the principal motive of effective communication. This seems to contradict a point Prys & Jones (2007) make earlier in the text, in a discussion

that seeks to rationalise the pursuit of lexical planning in the first place:

“Welsh speakers, in common with members of other small language communities, have a prejudice against borrowings from the adjacent dominant language and because of this, these informal borrowings are considered to be slang, and an effort is made to find native words, or to borrow from more acceptable languages, such as Latin”
(Prys & Jones, 2007: 7)

Here planning is presented as a solution to widespread borrowing, a solution that is considered necessary because borrowings, specifically from English, are viewed as ‘slang’, at least in registers that are not ‘colloquial and informal’. This comment seems to be made fundamentally within the realm of social meaning, and is at odds with Prys & Jones’ (2007) later insistence that the aim is to ‘promote effective communication’ and *not* to preserve linguistic purity. There seems to be a tension, then, between what drives lexical planning and how lexical planners rationalise their work²⁴.

That social meaning motives are implicated in lexical planning efforts is a notion that is addressed in the LPP literature. Fishman (1983, 2006) has written extensively about social aspects of term planning and lexical planning more broadly (although without necessarily referring explicitly to *social meaning*). He sees corpus planning as linguistic modernisation (making a language fit for use for modern purposes), but arguing that modernisation has to respond to the:

“particularistic directions into which and through which ‘universal modernization’ must be channelled. The amount of pull will vary.
The pullers and the pulled will vary. The point, however, remains

²⁴ I emphasise that social meaning relates to much more than attitudes towards borrowings and language planning. There are many issues relating to the contextualising of words, such as stylistic formality, interpersonal accommodation, lexical cohesion, etc. that impinge on lexical choice and usage (see section 2.5, below, on implementation).

valid: modernization drives, goals, needs, and processes alone are not enough for corpus planning to succeed. Modernization repeatedly needs to be particularistically digested, legitimated, and domesticated or disguised.” (Fishman, 1983: 111)

Words are coined, Fishman argues, with regard to these forces. For example, a common rationale for coining terms is to appeal to ‘folk’ or ‘man-in-the-street’ usage (1983: 113). A ‘folk’ form may have a certain perceived authenticity (from a certain ideological perspective), which might lend itself to acceptance by targeted users. However, as Fishman argues, from a different perspective ‘folk’ usage can be viewed negatively (e.g. as barbarisms, slang, etc.), and might not be considered by lexical planners.

What Fishman is referring to when he talks of ‘folk’ or ‘man-in-the-street’ usage is indexicality, a relationship of association between linguistic and social constructs. According to Ochs (2009: 406), “A linguistic index is usually a structure (...) that is used variably from one situation to another and becomes conventionally associated with particular situational dimensions such that when that structure is used, the form invokes those situational dimensions”. The situational dimensions that indices can refer to include, for example, social acts, social identities, epistemic stance, etc. In the case of purism, for example, a borrowing can index, through association with a category of people who are perceived as different, the social identity of the foreigner. However, indexicality doesn’t typically refer to a single situational dimension, but to a clustering of norms/expectations that are linked together through what Ochs calls indexical valences (Ochs, 2009: 412). In other words, certain indexical relationships entail others. For example, a borrowing, indexing the social identity of a foreigner, might also entail the threat that a foreigner might present. It is easy to see, then, in fraught language contact situations, how borrowing and other contact phenomena can come to be seen as symptomatic of ‘invasion’ and viewed in a highly negative light, something which is quite prevalent in some Welsh-language

discourses, as I have argued elsewhere (Robert, 2011).

Indexicality opens the door to understanding how lexical planning might come to be seen as motivated first and foremost by ‘effective communication’. I have argued that lexical planners conceive of their work as filling ‘lexical gaps’. But I have also argued that Welsh speakers can and do communicate in the absence of planning, largely by using borrowings and what is widely known in the bilingualism literature as ‘code-switching’. This is something Prys & Jones (2007) recognise, as we have seen, although they define the contexts in which such language is used as ‘colloquial and informal’. If linguistic forms, such as borrowing and ‘code-switches’ potentially index certain situational dimensions that are negatively evaluated, social identities in particular, it is not surprising that, within those situations where such an indexical relationship might be activated by borrowing or code-switching, speakers may not be willing to take on those identities. From this perspective, there *is* a lexical gap of sorts. It may not be a ‘gap’ that hinders the effective communication of propositional meaning, but rather one that hinders the communication of *social* meaning. It may be a ‘gap’ metalinguistically imagined, without any objective reality, but it is socially real. I will consider indexicality further later in the chapter, especially drawing attention to the ideological basis and ambiguity of indexical relations. But for the purposes of this section, what I have presented so far suffices to highlight how we might think differently about the aims of lexical planning projects, at least in the case of Welsh.

2.4 STANDARDISATION AND STANDARD LANGUAGES

In my discussion on what lexical planners mean when they describe their work as ‘standardisation’, I argued that what planners and term planning researchers call standardisation is probably more appropriately called codification, and that standardisation needs to be considered as impinging on social practice. What is standardisation, then? I want to continue by considering the nature of standardisation and standard language more generally – considering them as

theoretical constructs – drawing on the LPP and sociolinguistics literatures to investigate what it might mean to call a linguistic form or variety ‘standard’ and how this might come about.

Despite standardisation being a central concept in the LPP literature, language standardisation may or may not be the product of active, organised language planning. It seems more appropriate to consider lexical planning, or corpus planning more broadly, to be embedded within a wider culture of standardisation, a ‘drive’ towards uniformity, as J. Milroy (2002: 8) puts it. This ‘drive’ is similar to G. Lewis’ (1983: 309) depiction of language planning as part of a social ‘movement’, which he defines as “a series of actions which exemplify social tendencies and economic and political forces, as well as social and political principles or ideologies”. In the case of standardisation, seen as a movement, the tendency is towards uniformity in language practices and the belief in a unified language. This macro-perspective allows us to see corpus planning as a formal/organised manifestation of this broader movement, embedded in it and contributing to it at the same time. From this perspective, lexical planners in Wales carry out their work because of an already existing belief in the necessity of standardisation. In so doing, they make a contribution to and help consolidate the circulating discourse of standardisation. But this discourse is not lived only through the organised action of language planning. It also involves micro-level actions by individuals acting within (although not necessarily conforming to) their social milieu: spur-of-the-moment decisions regarding linguistic forms, conventionalised practices, corrective action (self- or other-targeted), etc.

This conceptualisation of standardisation as involving formalised activity (planning) as well as micro-level actions (language practices) chimes with the conceptualisation of language policy that is currently in vogue (McCarty, 2011; Schiffman, 1996; Shohamy, 2006; Spolsky, 2004; Wright, 2004): policy is seen not only as a formal declaration of proposed language-oriented action (the declaration that is the face of formal language planning), but also as a much more

fuzzy orientation to language (by an organisation or individual) that guides language practices, which speakers may not even be aware of. Schiffman lexicalises this distinction as *overt* and *covert* language policy, which he glosses respectively as “explicit, formalized, *de jure*, codified” and “implicit, informal, unstated, *de facto*, grass-roots” (1992: 3). But I find it is useful to differentiate between overt and covert on the one hand – relating to whether or not policy is conscious and explicitly motivated – and, on the other hand, formal and informal – relating to organisational principles (whether planning is officialised or ad hoc).

Standardisation, then, can be conceived of as involving the convergence of language practices within certain boundaries, maximally within all the language practices of a population (however that population is defined), resulting from collective language policies – formal or informal, overt or covert – that suppress, consciously or not, language variation. This seems to encompass both Haugen’s minimal variation in form, maximal variation in function (see above) as well as J. Milroy’s imposition of uniformity/suppression of variation. This definition, which sees standardisation as relating to language *practices*, also incorporates the attitudinal/ideological dimension. In Lewis’ definition above, where standardisation is seen as a social movement, standardising practice (what he calls ‘actions’) is seen to exemplify, inter alia, ‘social and political principles and ideologies’. Practice, then, has links to ideology, although the relationship between both is not always clear.

Language policy researchers have continued to emphasise the attitudinal/ideological dimension, arguing that, alongside practices and language management/planning, language policy resides in beliefs about language (Spolsky, 2004: 217), that is, in metalinguistic conceptualisations. Whether we want to conceive of metalinguistic conceptualisations as policy, as Spolsky seems to do, or as driving policy, is debatable. However, what I want to stress here is that standardisation is not all about the convergence of language practice, but

also has a metalinguistic dimension. This metalinguistic dimension is often referred to as the *ideology of standardisation* or *standard language ideology* (SLI) (Milroy & Milroy, 2012 [1985]).

What does standard language ideology entail? L. Milroy (1999: 174) notes that “The chief characteristic of a standard language ideology is the belief that there is one and only one correct spoken form of the language, modelled on a single correct written form”. Notwithstanding the fact that we’d do better to consider SLI to relate to all levels of language (Coupland, 2000), not only speech, L. Milroy’s definition relates both to a belief in the existence of language uniformity and to an evaluative dimension – that the ‘standard’ language is correct. In fact, SLI seems to relate to various evaluative stances. As J. Milroy notes (2001), *standard* means not only ‘commonplace’ or ‘normal’, but is used to describe a ‘measure of achievement’ – ‘good’ or ‘quality’. ‘Standard’ is a value-judgement. Moreover, ‘standard’ is relational: if ‘standard’ language is positively evaluated, other varieties become denigrated. This is how *slang* and *pidgin*, for example, get their designations. An early definition of French *patois* given in the *Encyclopédie*²⁵ is “Corrupt language as spoken in almost all the provinces... ‘Language’ proper is spoken only in the capital” (cited in Robb, 2007: 50). In this definition, *patois* is not merely bad language, it is so illegitimate it is hardly considered to be language at all. This is consistent with L. Milroy’s definition of SLI: if SLI entails a belief that a uniform language *exists* and that this uniform language is the only correct language, in the face of linguistic variation, ‘non-standard’ language must not be language at all. Standardness, then, relates to whether a linguistic form or variety is considered good or bad, proper or improper, normal or irregular, even language or not language.

But SLI entails more than beliefs about language. Coupland (2000: 627) argues, in relation to ‘standard English’, that SLI relates to the “social positioning that

²⁵ I assume Robb means Diderot & d’Alembert’s *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers*, which was published in the latter half of the 18th Century.

goes on around the notion of language standardness". That is, the ideology of the standard doesn't just refer to some abstract beliefs that exist in a separate dimension, but it is manifested and lived through the ways that speakers relate to each other and construct social relations through language practices. This brings us back to practice, but not practice as value-free, apolitical use of linguistic forms, but language practice shaping the social world and managing social relations.

Here I would like to return to the notion of indexicality that I discussed earlier – a relationship of association between a linguistic form and a social construct, such as a social identity, situation, relationship, etc. such that, when the linguistic form is used, the social construct is evoked, very simply put. If 'standard' forms are considered 'normal' or 'quality' (or even 'standard'), we can consider there to be an indexical relationship between linguistic forms that are considered 'standard' and the qualities that are attributed to them. In fact, the word 'standard' itself suggests an indexical relationship. There is nothing inherently 'standard' about any linguistic form: the label 'standard' represents a normative assumption, a social construct, not an objective reality. This normative assumption, however, positions 'standard' language as meaning-free.

This is demonstrated in dictionaries and glossaries. As Bourdieu (1991: 48) points out, dictionaries go to the lengths of "marking those [words] which exceed the bounds of acceptability with a sign of exclusion such as *Obs.*, *Coll.* or *Sl.*", which singles those words out as non-standard/-legitimate. However, other words are not marked at all. That is, they are presented as *normal*. (Note that these words are also usually the first words listed under a dictionary entry, that is, they are given priority.) The social meaning that these words carry (or that planners would have them carry) is rendered invisible, and they are presented as value-free. In the social positioning going on here, 'standard' forms are assumed to do no social positioning work. They are positioned as *normalised*. That is, they are positioned as being "capable of functioning outside the constraints and

without the assistance of the situation, and [...] suitable for transmitting and decoding by any sender and receiver, who may know nothing of one another” (Bourdieu, 1991: 48). They index nothing. However, as Bourdieu (1991) insists, language acquisition and use are always socially and economically conditioned, hence there is no value-free language. It is a point also made in the sociolinguistic literature on indexicality: how linguistic forms signify will shift depending on perspective (Coupland, 2007; Ochs, 2009; Woolard, 2005). Perspective brings us back to standard language ideology.

The means of normalisation and legitimation (how a linguistic form or variety comes to be perceived as ‘standard’) can be explained by looking in more detail at the concept of language ideology. In their afterword to the 2012 edition of *Authority in Language*, Milroy & Milroy (2012 [1985]) reconsider standard language ideology in light of the research on language ideologies that had been written in the years since the book was first published. Drawing in particular on the work of Michael Silverstein and Judith Irvine, they make the point that ideologies are cultural systems of beliefs that are related to social and political interests. From this perspective, they infer that everybody is influenced by a range of dominant ideologies. However, since ideologies are related to social and political interests, it is the ideologies of elites that become prominent and gain most credibility.

Nonetheless, there is still the question of how this comes to bear, and this is an important question for my research since it can shed light on the likely impact of Welsh lexical planning. Anderson’s (2006 [1983]) account of the development of national standard languages relates directly to the printing press and the role it had in imposing uniformity in written language for technical and economic-ideological reasons at the same time as changing social consciousness of community and language community (what he calls ‘imagined community’). That is, the capitalist printing endeavour and its effects resulted in a change in people’s ideas about who they were and who they shared a common language

with, and hence resulted in the definition of national standard languages. Note that Anderson's account links standard languages directly to technology and economy as well as to politics, namely nationalism. Although Anderson's account relates to the *emergence* of national standard languages and standard language ideology, we can see how it might be relevant to their continuation, although by the same token, more recent technological, political and economic developments might be changing consciousness of language community membership.

Bourdieu (1991) has offered a somewhat different account, which emphasises the role of power and authority, in particular through state institutions. Bourdieu argues that the process of legitimisation is brought about through the

“dialectical relation between the school system and the labour market – or, more precisely, between the unification of the educational (and linguistic) market, linked to the introduction of educational qualifications valid nation-wide, independent (at least officially) of the social or regional characteristics of their bearers, and the unification of the labour market (including the development of the state administration and the civil service).” (1991: 49)

That is, in order to be valued on the labour market, a market which has been unified through state centralisation, one needs to gain qualifications accredited by the state education system, the system which, more than any other, endorses the standard language ideology.

Without wanting to claim that this fully accounts for the process of standardisation, we can see how the ideologies of socially and economically powerful groups, in Milroy & Milroy's understanding of ideology, might become prominent – through their acquisition in education institutions and their enforcement in the labour market. This can also account for the hegemony of the standard language ideology where speakers who have limited proficiency in the

‘standard’ language, to the extent that standard language is locatable, orient to the ‘standard’ as legitimate language. Woolard (1985: 741) argues that this is, in fact, the typical situation in a linguistic culture of standardisation – convergence of metalinguistic conceptualisation without total convergence of language practices.

Since, according to this conceptualisation of standardisation, power and authority are central to standardisation and SLI, they are only as powerful as the institutions that endorse them. There are always competing ideologies. Milroy & Milroy talk of ‘vernacular maintenance’ (1997: 53) or ‘covert and informal language maintenance’ (2012 [1985]: 49), which is “exerted by members of one’s peer-group or social group”, and not endorsed by the institutions of the state. When I talked about a social movement, then, at the beginning of the section, there isn’t necessarily a uniform drive towards uniformity, but more of a tug-of-war or an ideological contest (Coupland, 2007: 42): a movement amongst other movements competing for social space. This brings us back to the subjective nature of indexicality. Whether a linguistic form or variety is considered to be ‘standard’ – whether it is considered normal, suitable for use with anybody anywhere – relates to a point of view in relation to that contest (Coupland, 2007: 42). This is one of the reasons that Coupland dislikes the ascription ‘standard’ –

“it forecloses on ideological conflict and on its outcomes. It presupposes that there is a set of linguistic forms whose social value is known and uniform” (2007: 43)

What conclusions can we draw from this discussion in relation to lexical planning for Welsh? For their planned words to become ‘standard’, lexical planners are aiming for normalisation, that is, that planned words carry neutral social meaning. This ought to be true whether planners are planning for a unified standard language or for an elaborated technical register. They seek to achieve this partly through their marking conventions in codification texts – positioning

words as 'normal'. But this positioning work does not necessarily secure absolute normalisation. Planners' efforts to legitimise their own planned language do not necessarily lead to speakers orienting to the planned words in the same way. Words in use might achieve very different social meanings. In terms of the implementation of lexical planning, the argument against merely asking *whether* planned words are used becomes more clearly focused. We need to go further than this in order to consider the take-up of lexical planning. Even if planned words are used, do they mean what planners intend for them to mean? What social positioning work do planned words do in interaction? This is an important motivation behind my empirical analyses, which will both 'look backwards' at the institutional credentials of particular terms, and 'look inwards' at the local contextualisations of terms in their own discursive contexts.

The literature on the emergence and consolidation of standard languages suggests that the full standardisation of any minority language would be difficult to achieve. Although there has emerged some form of national consciousness where language is implicated, partly imagined, as we might expect according to Anderson's account, through literature, poetry, journalism²⁶, etc. this national consciousness is not fully endorsed by national institutions. If, according to Bourdieu's account, a standard language is only as powerful as the institutions that endorse it, this is something of a Catch-22 for minority languages, since they are minoritised as a result of the power of those very same institutions. Although (as I will explain in Chapter 4) there is widespread Welsh-medium education in Wales (up to 18 years), for example, there is very little Welsh-medium higher education or, perhaps more significantly in terms of Bourdieu's account, professional training. If standardisation is contingent on the integration of the labour market with a unified education system, Welsh isn't likely to be thoroughly standardised in the foreseeable future. Legislative moves in recent decades have led to Welsh-language skills being a requirement for a number of

²⁶ Welsh-language newspapers had their heyday in the 19th century, although there has never been a daily Welsh-language newspaper. Currently, there is only one Welsh-language 'national' newspaper, the weekly *Y Cymro*. See ap Dyfrig et al. (2006) for more details.

jobs, but, with the exception of certain domains (notably media and education), it is unclear how far these jobs go beyond language professions and language-specialised roles in public services. It is also unclear, in the case of the latter, whether Welsh-speaking employees are embedded within a workplace culture where Welsh is the working language. Moreover, as we shall see in Chapter 4, corpus planning seems to be relatively low on the Government's language-planning agenda. It might be argued that the political climate in Wales is more conducive to standardisation initiatives than at any time in the past, considering the increased use of Welsh in media, education and public services. However, currently Welsh seems to lack the political clout to be standardised thoroughly (in as far as thorough standardisation is possible at all). For this to happen, according to Bourdieu's theory, (standard) Welsh would have to be initially defined far more stringently than is currently the case, then to become a required skill for the vast majority of jobs (or perhaps select, powerful jobs), as well as be taught thoroughly across the vast majority of schools in Wales, something that would need a great deal of political will to drive through and would likely alienate the majority of voters. What this means, then, for my research is that I do not expect to find lexical planning thoroughly implemented.

2.5 IMPLEMENTATION

One of the main concerns of this thesis is the *implementation* of lexical planning. It has frequently been noted that language planning cannot be considered successful unless it is implemented, although implementation is often considered to be synonymous with acceptance. For example, Fishman (1983: 109) argues that the acceptance of language planning, that is to have the language "liked, learned and used", must be the goal of language planning. Others who have expressed similar arguments include Mwansoko (1993), Bhreathnach (2012) and Jernudd & Das Gupta (1971). This opinion is also reflected in the Welsh Language Board's desire to plan language that is "popular, useful and used" (Welsh Language Board, 2005: 36). However, there is potentially more to implementation than a willing acceptance of planned language by its target users,

as I argue below.

Implementation is one of the four stages of Haugen's (1983) model of language planning – code selection, codification, implementation and elaboration.

Although Haugen wrote about the implementation stage in his model, he didn't develop a theory of what implementation entails. By his own admission, the topic needed a researcher with expertise different from his own. However, Haugen does define implementation as "the activity of a writer, an institution, a government in *adopting* and *attempting to spread* the language form that has been selected and codified" (1983: 272, my emphasis). Term planning researchers are also concerned with implementation, formally differentiating between two separate stages, namely *dissemination* and *implantation* (e.g. Bhreathnach, 2012). These stages are analogous with Haugen's 'attempt to spread' (dissemination) the planned language and its being adopted (implantation). However, I find it useful to differentiate instead between *implementation measures* and the *take-up* of planned language in linguistic practice. Implementation measures can be conceived of as deliberate, although not always organised, attempts by planners or policy-makers to effect the acquisition and use of their planned language, what Coupland & Kristiansen (2011: 22) refer to as "top-down, controlling activities by national governments and their agencies". They include enforcement, persuasion and dissemination measures, although sometimes the difference between these three isn't clear. It can also be difficult to differentiate between implementation measures and the take-up of planned language in linguistic practice. That is, micro language practices (the use of the planned language) may result – deliberately or not – in dissemination and result in others using the language.

An example of a corpus-planning focused implementation measure is found in the French state's Bas-Lauriol law of 1975. Weinstein (1989) writes that the law decreed that 'the French language' must be used in any communication that dealt with the sale of goods and services, as well as all labour contracts, signs on public

buildings and parties receiving public funds. ‘The French language’ was defined in puristic terms – no foreign words or expressions were allowed, at least not when a French-language equivalent was available, as listed in the glossary of the *Haute Comité pour la Défense et l’Expansion de la Langue Française*. Infractions were punishable as fraud. It was, then, a measure that simultaneously enforced the use of planned terms and prohibited borrowings, construing the latter as a deliberate attempt at deception²⁷.

French laws relating to corpus policy have been difficult to enforce, however, partly because of the leniency of punishment (Cartrite, 2009; Weinstein, 1989). Despite being punishable as fraud, the fines incurred as punishment were small. Weinstein also notes that the Bas-Lauriol law was not widely-known, and he is sceptical whether most French people consult prescriptive dictionaries, or whether the dictionaries most widely used are as strict in their purist orientation.

The above highlights that implementation measures do not necessarily lead to the take-up of planned language at the micro level of language practices, even in France! Language policy and planning researchers have highlighted that the implementation of language policy is dependent on and carried out through the micro-level actions of individuals within the social contexts in which they act (e.g. Schiffman, 1992). This perspective shifts the focus “from decisions made at the top of the institutional or status hierarchy to acceptance obtained from the speech community in general” (Coupland & Kristiansen, 2011: 22). For example, language-in-education policy is partly implemented through the actions of teachers within their institutional settings. Government policies may specify how the implementation of the policy is to be gauged and evaluated, and what action should be taken if the policy is not implemented. Yet research in, for example,

²⁷ In 1994, the Bas-Lauriol law was replaced by the Toubon law. The new law made similar provisions as regards term planning and borrowing, but it was designed to be more rigorously enforced whilst making some provisions to safeguard freedom of speech (Cartrite, 2009; Oakes, 2001).

education settings shows a more mixed picture, where a policy may be implemented for some tasks or actions and not for others, sometimes as a result of conflicting demands. Based on research she conducted with colleagues on bilingual education in north-west England (cited in Martin-Jones, 1995), Martin-Jones inferred that bilingual classroom assistants, who had officially been appointed to support those whose home language was not English to learn the language, did not rigidly carry out the institutional priorities in the classroom, but varied their language practices according to their interactional goals. Specifically, in talk with learners that focused on the curriculum content, they typically started with a mix of English and the home language (Punjabi in the example cited) and elicited responses in English. But they then switched to the home language for more specific instructions in order to develop their care-giving, close relationships with the learners.

This research perspective emphasises the *agency* of individual actors. Such research sees individuals as social actors having the potential, and indeed the obligation, to make their own choices regarding their language practices within their social contexts, rather than blindly adhering to “top-down controlling activities by national governments and their agencies” (Coupland & Kristiansen, 2011: 22). ‘Choices’ are not necessarily consciously made or open to reflection by social actors, but the view of speakers as social actors positions them as using language in orienting to and participating in their social worlds, and having some degree of agency in how they do so. To return to the idea of standardisation and social meaning, and to reiterate the question I set out there, where language users do use planned lexis, does it achieve the same social meaning that planners intend? The flipside of this question is, where language users use lexis that is not planned, is this done within the same ideological framework that planners take to language? In other words, do speakers use non-legitimised language as non-legitimate, or does it index other social relations and identities?

Viewing language practice as social action does not entail, however, that speakers

are completely free to make any choices they wish. Coupland (2007) emphasises, for example, that social actors do not necessarily have full control over their linguistic behaviour. In part, this is due to the deeply in-grained, habitualised nature of language and the competencies of individual speakers (a point Cameron (2009 [1990]) also makes). Other constraints may be related to social and institutional context, where 'institutional' is defined in relation to social legitimation through structures of authority (Mayr, 2008). As we have already seen in the discussion on standardisation, 'standard languages' are powerful, state-authorised symbolic codes. In what contexts do language users feel constrained or obligated to use this symbolic code, and why? We also saw in the previous section, however, that standardisation is an ideological contest. What other norms and ideologies are available to speakers? When and why does 'vernacular maintenance' apply? Do we find conflict between ideologies within the same situation? In the Martin-Jones example, perhaps the bilingual assistants felt obligated by the education institution to implement official policy through their curriculum-oriented talk. But what compelled them to nurture a close relationship with the learners through their learner-oriented talk? Their language practices suggested a negotiation of conflicting demands and goals. Yet other constraints might involve pragmatic considerations of the communication of propositional meaning. In particular in the case of lexical planning, how do language users orient to the use of neologisms that may not be familiar to their audience?

Implementation, then, isn't a simple case of the language planned being used or not. Rather, considering the take-up of lexical planning requires an approach that prioritises investigating language practices, and views the use (or not) of planned language as social action carried out by social actors whose language choices are contingent on a number of factors. These factors include how social actors try to mean, their language faculties, and institutional and ideological as well as pragmatic constraints. Such an approach places the research squarely in the realm of sociolinguistics and of the critical analysis of discourse. We can

formulate some general research questions from these conclusions: Do language users use planned lexis? What social meanings are activated by the use of planned lexis? And what constrains the use of planned lexis in interaction?

CHAPTER THREE:

Methods and data

3.1 INTRODUCTION

The research questions I have laid out on page 45 prioritise the study of language practices for considering the implementation and social embedding of lexical planning. My approach investigates language practices through analysing recordings at two data sites. I approach the data from two different perspectives. First, I offer an account of the degree to which participants take up planned words in the data. This approach considers the degree to which lexical items used by participants in the data coincide with those prescribed by planners. It presents language as a decontextualised resource, involving selecting a sample of lexical items from the data and comparing them with the recommendations/stipulations of selected codification texts. This approach goes some way to answering whether participants use the planned terms in their language practices. In its quantitative approach it bears some relation to corpus linguistics and to variationist sociolinguistics, where selected linguistic forms are annotated or coded across a body of texts and collated to give a quantitative account of the distribution of those forms across that body of texts. Second, I present an interactional analysis of participants' choice of words, inquiring into how and why participants use the words they do. In Chapter 1, I argued that an investigation of the take-up of corpus planning needs to consider whether and how planned language gets embedded in its social milieu. Whether it is used gives some indication as to whether it is picked up, but it doesn't tell us much about whether speakers feel ownership over it, whether it facilitates the use of Welsh, whether it is universally helpful or sometimes problematic. An interactional analysis of in situ language policy and practice is conducted, then, to investigate these questions, by considering how and whether planned language is used by participants to achieve their interactional goals (although see my discussion on what I mean by 'goals' in section 3.6). This approach draws on

more socially-oriented language research, such as linguistic ethnography, treating “the interface between language/text and situation/context as a central problem” (Rampton, 2007: 589). It treats the language that surfaces in the data as embedded within its social context rather than separated from it.

What language practices and social environments can we investigate to consider the take-up of lexical planning? The focus of recent lexical planning work in Wales on subject-specific vocabulary makes it pertinent to look for talk that deals with some of the areas that have been planned for. Although I have argued that lexical planning is broader than Terminology or LSP, and that lexicographers (as opposed to terminologists) engage in lexical planning without a view on subject specificity per se, undoubtedly much of the work that has been carried out to date focuses on specific subject areas. Therefore, whilst I concentrate on talk where vocabulary might be focused on particular subjects, it is to be borne in mind that this sits within a broader lexical planning effort.

Another reason for researching language practices where vocabulary might be focused is my previous argument that lexical planning is carried out with an eye on status planning for language revitalisation. That is, corpus planning is largely carried out to aid the domain-extension of Welsh. G. Williams argues that ‘domains’ – the target of status planning as the theoretical sites in which function (re)allocation is envisaged – need to be understood to “constitute *institutional* settings which are vested with power, authority, and control. As such, they constitute the agencies for the reproduction of languages” (Williams, 1987: 86, my emphasis). If language planning is concerned with the (re)production of Welsh, as I think it must be (since ‘(re)production’ is another way of referring to the *doing* of language), then it makes sense to consider how language is used in those domains/institutions that are most crucial for social (re)production. Two of those domains in (late-)modern society are arguably the broadcast media and the workplace. Even if it is unclear quite what is the effect of revitalisation efforts that focus on these institutions, as in the case of the media (Cormack, 2007),

certainly in the case of Welsh, workplaces and the media have received a great deal of attention by planners and language activists (I go into more detail in the next chapter). Moreover, it has been argued in the case of the media (Busch, 2010; Williams, 1992: 141) and the workplace (if we take Bourdieu's (1991: 49) 'labour markets' to be directly related to workplaces) that they are central to the legitimisation of the 'standard' language. Therefore, they are very relevant to an investigation of the implementation of planning.

In the rest of this chapter I describe the data collection methods and introduce the datasets themselves, before outlining in more detail how I set out to analyse the data and why I made the choices I did. First, however, I consider the ethical considerations and implications of the research.

3.2 ETHICAL CONSIDERATIONS

Both datasets are very different in nature and in the way they were collected. Both involve naturally-occurring speech events, but the workplace data was recorded *as data*, with the explicit, informed consent of the participants, whilst the media data was recorded as a live broadcast. This meant that participants in the media data did not give their informed consent for the collection or analysis of the data. However, since the broadcasts had been uploaded to the show's website, as well as later made available to me as digital files, I felt free to use them as data. The anonymisation of media texts may not be necessary, or even possible, considering the public nature of the data.

Nonetheless, this does not relieve me from any ethical obligation to the participants, since there are ethical considerations beyond anonymity. Duranti (1997: 120-1) argues that:

"There is [...] no way of escaping the responsibility we have as researchers towards the people we study. This does not mean that we should always and only write what we think they will like, but

that whatever we decide to say publicly and publish should be informed by our awareness of the potential consequences of our research.”

I consider that this applies to my media data as much as to any other data, in this case specifically relating to representation – what we say about participants and how we say it in the presentation of research. It means taking care over how I present participants, ensuring that my interpretations are fair and balanced, in order to “[show] due respect to all participants [and] to the values of truth, fairness and open democracy” (The British Association for Applied Linguistics, 2006: 2). As well as taking care with representation, I have taken some steps in anonymising the media data by not naming guests in the presentation of my research. Instead, I refer to ‘the guest in [text name]’. Although the texts are of a public nature, concealing guests’ names in my own work introduces a level of anonymity, where readers would have to take active steps to find their names. It also prevents guest names from being made available in digitised text, which might be retrievable, for example, by an online search engine.

There are more complex ethical issues involved with the spontaneous speech recordings of the workplace data. I am ethically accountable to the individuals whose speech acts make up the recorded data as well as to the institution in which the dataset is embedded, and who enabled my research by allowing me to record at their site. Being ethically accountable includes the legal obligations stipulated by the UK Data Protection Act 1998 (see the legal guidance to the Act in: Information Commissioner's Office, n.d.), but extends to the moral obligation of ensuring that I “respect the rights, interests, sensitivities, and privacy” of participants (The British Association for Applied Linguistics, 2006: 4).

Before making any recordings I gained the explicit informed consent of all participants – individual participants and the organisation itself (see Appendix 1). This involved establishing rights of access to the data and making clear to

participants that participation was voluntary. However, researchers increasingly recognise that informed consent is a slippery issue (The British Association for Applied Linguistics, 2006: 4). Whether participants give their full informed consent voluntarily depends, for example, on: participants' understanding of the nature of academic analysis and publication; their understanding of the information sheet text (is it written clearly, in plain language? can they read?); whether they feel pressurised (by the researcher or others) to give consent; and whether they have the time or inclination to read the information sheet with care, if at all.

For these reasons, whilst seeking the voluntary, explicit, informed consent of participants is paramount, consent must also be balanced by continuously being aware of the potential implications of the research for participants. This involves careful consideration of how to anonymise the data and an awareness of the sensitivity of the data collected. Both of these points need to be considered in the processing (e.g. transcription) and storing of the data, but particularly in deciding which content to make available in presenting the research (e.g. in the form of this thesis).

With this in mind, I have sought to anonymise the data to the best of my ability. Participants have been assigned pseudonyms. Pseudonyms have not been selected for their social equivalence (Rock, 2001) with the original names, other than gender, lest this practice should reveal too much of a participant's identity. The Housing Association is not named. It is always referred to as 'the (Housing) Association' or 'the organisation'. Its location is only given as north-west Wales, and its projects and properties have either been given pseudonyms or have been replaced by '(house name)', '(street name)', etc²⁸. Rock (2001) reminds us that we also have a duty to third-parties who might be named by participants in the data, in particular if the third-party is presented in a negative light. Third-parties,

²⁸ Since it is a housing association, it holds a number of properties, which are often referred to in the recordings. I have typically given the properties most salient in the data pseudonyms and use the other method for lesser-mentioned properties.

then, have also been given pseudonyms.

However, anonymisation requires concealing participants' identities as well as their names. Identity may be revealed by any number of factors: linguistic features such as accent or idiosyncratic turns of phrase; paralinguistic features, such as voice; formal labels, such as job title; or even the content of talk, such as a story relating to a particular individual. Other names or titles mentioned, then, such as organisations, job titles, place names, etc. by which the participants might be identified have been replaced by '(job title)', '(company name)', etc. Moreover, in the presentation of my research, I try not to present data that might reveal details by which participants can be recognised. However, I should note that the possibility of identification varies according to who does the perceiving and how closely acquainted they are with the participant in question. Rock (2001: 15) differentiates between 'insiders' and 'outsiders', arguing that "insiders should be protected from being recognised by other insiders however unlikely that scenario may seem". In research relating to Welsh, the pool of Welsh speakers from which to draw participants is small, and the probability that somebody reading this work is an 'insider' to the participants is high. Hornsby (2009) has made a similar observation in relation to his research with neo-Breton speakers. When we also account for the need, for the interpretation of the data, to identify the nature of the work of the organisation, its broad location, the fact that it works through the medium of Welsh, as well as some information relating to the participants' role within the organisation (the nature of their jobs or their position within the organisational hierarchy²⁹), the matter of anonymisation becomes difficult.

It is not unusual for anonymisation to cause some difficulty, since sociolinguistic researchers "often face a variety of conflicting interests and competing obligations" (The British Association for Applied Linguistics, 2006: 2). As Rock

²⁹ Job titles are an example of information about participants I would have liked to include, but decided against for reasons of anonymisation. They are interesting from a lexical planning perspective because they are institutionally-vetted, fixed labels.

(2001: 2) argues, “anonymisation necessitates locating one's research on a balance weighted on one side by a need to render data subjects totally unrecognisable and on the other by a desire not to remove information from data”. Again, we can turn to Duranti's (1997) words (cited on p.49-50) that what is important is to be aware of the potential consequences of our research, and not any uncritical, generalised approach to anonymisation. This requires careful attention to the sensitivity of the data. I have taken care to delete sections from the data that participants asked to be deleted. I take care not to present any content that might be considered sensitive and, as with the media dataset, I take care in the way that I represent the individual participants and the organisation, ensuring fair and balanced interpretation. Whilst I have made every effort to anonymise the data, considering the difficulty of the task I hope that my careful treatment of the data ensures that my needs as a researcher are balanced with the interests of my participants.

The above discussion deals with most of the ethical considerations relating to the data, but others are considered within the specific context of data collection in the next two sections.

3.3 MEDIA DATA: DR ANN'S *SYRJERI*

The data collected for the media study is primarily a collection of video texts taken from *Wedi3* ('After 3'), a day-time magazine show which is broadcast on the Welsh-language television channel, S4C (*Sianel Pedwar Cymru*, or 'Channel Four Wales') every weekday at 3pm³⁰. *Wedi3* is made for S4C by production company *Tinopolis* in its studios in Llanelli, south-west Wales. *Wedi3* is broadcast live. Each broadcast consists of a number of what I call *items*. Items focus on particular themes or activities, such as advice on gardening, a cooking demonstration, advice on fashion, etc., which make *Wedi3* fertile ground for data involving subject-specific lexis.

³⁰ *Wedi3* and its sister programme, *Wedi7*, ceased broadcasting in February 2012.

One of the regular items on *Wedi3* is *Y Syrjeri* ('the Surgery'). This is an item primarily broadcast live from the studio, involving the presenter(s) for the day, Dr Ann (a health expert) and occasionally a guest. The participants discuss a particular medical condition or a health-related issue. Conditions or issues include, for example, diabetes, the kidneys, and children's health. The health-focus of the *Syrjeri* items met the need of subject-focused talk, in particular in a field where a lot of lexical planning had been carried out. On the *Wedi3* website³¹, where programme content is uploaded as individual items after each broadcast, each *Syrjeri* item is named according to the medical condition or issue that is being discussed. The data used for the study consists of twenty *Syrjeri* items (specifically the electronic files of individual items created for the online archive) that were broadcast between October 2008 and October 2009, and corresponding transcripts. The air time of all items adds up to approximately 2 hours 28 minutes.

From the audiovisual data, I selected the items for analysis from amongst the numerous *Syrjeri* items that were available on the *Wedi3* website³². The items used are as follows (labels are taken from the Welsh and English versions of the website):

<i>Bynions a Chyrn</i>	<i>Bunions and Corns</i>
<i>Cerrig y Bustl</i>	<i>Gallstones</i>
<i>Clefyd y Siwgr</i>	<i>Diabetes</i>
<i>Colli Gwallt</i>	<i>Alopecia</i>
<i>Ffliw Moch a Dirywiad</i>	<i>Swine Flu and Macular</i>
<i>Macwlaidd</i>	<i>Degeneration</i>
<i>Eczema</i>	<i>Eczema</i>
<i>Epilepsi</i>	<i>Epilepsy</i>
<i>Ffliw'r Moch(1)</i>	<i>Swine Flu(1)</i>
<i>Ffliw'r Moch(2)</i> ³³	<i>Swine Flu(2)</i>

³¹ www.wedi7.com (*Wedi3* shares a website with its sister programme, *Wedi7*).

³² *Syrjeri* items are not labelled as such on the website. However, all the items selected are identifiable as *Syrjeri* items since they tend to follow a similar pattern of discourse, which includes similar content and participants. In addition, a number of the items are introduced by presenters during the broadcast as "y syrjeri", and they were all archived under 'Health and Wellbeing' on the website.

³³ This text wasn't available on the *Wedi3* website, but was sent to me with the other texts when I inquired after the files.

<i>Y Frech Goch</i>	<i>Measles</i>
<i>Iechyd Plant</i>	<i>Children's Health</i>
<i>Iechyd y Galon</i>	<i>Cardiac Health</i>
<i>Iselder</i>	<i>Depression</i>
<i>Lwpws</i>	<i>Lupus</i>
<i>Marwolaeth y Crud</i>	<i>Cot Death</i>
<i>Llid yr Ymennydd</i>	<i>Meningitis</i>
<i>Sglerosis Ymledol</i>	<i>Multiple Sclerosis</i>
<i>Osteoporosis</i>	<i>Osteoporosis</i>
<i>Pendics</i>	<i>Appendix</i>
<i>Y Dwymyn Goch</i>	<i>Scarlet Fever</i>

Henceforth, I use the English label (italicised) to refer to the items as broadcast (that is, the video files, which are the 'texts'). I refer to their respective transcripts when specified.

In addition to the audiovisual data, I spent a day at the *Wedi3* studios, where I observed the programme being produced live. I also got a chance to talk to some of the show's researchers and to Dr Ann, and sat in on a production meeting, which involved the two presenters for the day. I wrote ethnographic-type notes of the day, which I use to aid interpretation. However, my analysis focuses first and foremost on the audiovisual data of the television broadcasts.

S4C also has a language policy, its *Language Guidelines* (S4C, 2008), which I have drawn on to aid the interpretation of the data. The guidelines are used as a means of offering a picture of the wider institutional context in which the interactional data is embedded. This wider institutional context is part of what Fairclough (1995b: 57) calls the 'socio-cultural goings on' of the data – the background to the data, as well as the context of which the data, as communicative event is part. In Fairclough's critical discourse approach, the discussion of the policy document would be treated as part of the analysis. His three-pronged approach to analysis involves text analysis (e.g. the conversational structure), analysis of discourse practices (the way texts are produced, received and distributed) and analysis of sociocultural practices (on situational, institutional and societal levels). The policy analysis, as analysis of sociocultural

practices, contributes to answering the *why* of lexical choice. It would therefore sit comfortably within an analysis chapter. On the other hand, there are other aspects to the ‘socio-cultural goings on’ of the data, which lie outside of the media institution. These include a consideration of sociolinguistic variation and standardisation. I deal with these in the next chapter, which sets out the sociolinguistic context to the data. Since the policy document makes the same type of contribution to the analysis, I consider S4C’s language policy within the same framework, treating it, then, as aiding interpretation of the other data, rather than data in and of itself.

3.4 WORKPLACE DATA: HOUSING ASSOCIATION

The workplace data was collected at a housing association³⁴ located in north-west Wales. Locating a workplace in which to base the research posed some problems. Following the 1993 Welsh Language Act, a great many organisations are under some obligation to use Welsh in their dealings with the public. However, with the exception of schools and translation firms, which are workplaces of a different nature than what I had in mind, there are few organisations where Welsh is a language of daily interaction, the primary criterion for selecting a site for my study. Moreover, collecting data at a workplace, or any organisation, requires gaining the consent of the managers of the organisation. Finding a workplace, then, where Welsh was the/a language of daily communication, where gatekeepers would be happy to let in a researcher with equipment to record potentially sensitive data was a challenge, and the choice of workplace was largely driven by these concerns.

The interactional data collected consists of recordings of four meetings held at the Association’s premises. Meetings were chosen as data primarily for logistical reasons: they are mostly closed events, in that they involve little coming and going of participants, which makes gaining informed consent easier and reduces

³⁴ Housing associations are not-for-profit organisations that provide social housing (low cost housing for people in most need).

the risk of recording people without their knowing; they are less intrusive on the organisation and its staff; the researcher can be present at the speech events with minimal influence on the interaction, since participants are primarily concerned with getting on with the business of the meeting; and they were the types of event the organisation was happy for me to record as data. In terms of the content of the data, meetings were a good choice for this research because: they cover a lot of different topics, typically bringing up a number of lexical items commonly used within the organisation; they are talk intensive (compared with desk work, for example); and they typically involve different types of interaction and interactional goals, such as where participants strive to perform professional identities, or where they nurture personal relationships and camaraderie with their colleagues.

I was present throughout the recordings, sitting at the table with the participants or nearby. I had arranged with my contact at the Association in advance which meetings I would attend. I had asked for information sheets and consent forms to be sent to participants by email beforehand, and I collected them at the beginning of each meeting, after being introduced. I would set up and start the recording and the meeting would proceed mostly as normal. On the whole, it doesn't seem that my presence affected the interaction very significantly. Considering the need for participants to concentrate on the interaction (especially those who did most of the talking) and the length of the meetings, it seemed that it should be easy for participants to forget about being recorded, or at least not to pay it too much heed. There are very few instances where I am referred to explicitly in the data.

The four meetings recorded cover different topics and all involve different participants, although there is considerable overlap. This has given me a potential dataset considerably larger than the media dataset, coming to more than 6 hours and 34 participants, the vast majority of which make spoken contributions. I chose to restrict the analysis to certain sections of the four

recordings, amounting to 2 hours 30 minutes of recorded time, to make the dataset used for analysis comparable in length to the media dataset³⁵.

I have given each recording a title, which is a shorthand translation of the title given the original meeting. *Voids* was a meeting of the *Voids Working Group*³⁶, who met to discuss vacant properties on the Association's list and how to find tenants for them. *Estates* was a meeting held by the *Tenancies Team* to discuss visits that the Team's staff had made to some of their estates to gather information about problems with the properties or their management that need consideration. *Minafon* was a meeting of the *Minafon Project Team* who are responsible for overseeing a property development at the Association's *Minafon* estate ('Minafon' is a pseudonym). Finally, *Publicaitions* was a meeting of the *Publicaitions Panel*, which consists of tenants and some of the Association's staff, who discuss the publications the Association gives its tenants as formal channels of communication, such as the tenants' handbook.

Like in the case of the media data, I gathered some amount of what might be called auxiliary data, which I refer to in the analysis, when relevant, to support my interpretation of the primary data, the audio recordings. These include, for example, agendas, minutes of the previous meeting, project budgets or an example of a tenants' handbook used for group discussion. At some meetings, visual materials were projected from the laptop onto the screen, and in one meeting notes were written on the flipchart. Often the documentation and visual materials are helpful in deciphering what it is participants are talking about in the recordings, since they form part of the interaction, where participants refer to them. But at times, these materials are interesting for their language content that is specifically relevant to the subject-matter of my research. Other auxiliary data includes what I observed from being present at the meetings on the Association's premises, and from my contact with the staff.

³⁵ From this selection, I have also selected a sub-sample for the quantitative analysis (see below).

³⁶ The organisation used Welsh titles for all these teams, working groups, etc., but I have given them English translations here.

Finally, I have referred to some documentation from outside of the immediate context of the meetings in order to consider the wider institutional framework in which the data is embedded. In this, I seek to follow a similar approach to the one taken for the media study, where I discuss S4C's formal language policy document in the next chapter, as part of the wider sociolinguistic context of the data. Although the discussion is used to aid data analysis, it is not treated as analysis in and of itself. The Housing Association doesn't have a comparable formal language policy, but some inferences can be made regarding language evaluation and ideology on an institutional level from the documentation available, in particular its website and Welsh Language Scheme. Note, however, that I do not name this documentation or quote directly from it, to try to protect the identities of the participants and the organisation.

3.5 SUBJECT-SPECIFIC LEXIS AS CODIFIED

This approach seeks to give an impression of the extent to which the subject-specific lexis found in the data can be said to be 'codified', or to be in harmony with the stipulations made in lexical codification texts. It takes the vocabulary found in the data that can be considered subject-specific or work-related, such as *gwrthbiotig* ('antibiotic') or *ebost* ('email'), comparing it with specified codification texts. I categorise the words selected according to whether or not they can be considered codified, according to the texts consulted, and give the distribution of these categories, rather than of the individual words. It is, after all, the fact of their being codified or not that is of most interest in this analytic approach, rather than each word in and of itself.

There were a number of factors to be considered in designing and conducting this part of the analysis, such as how much of the data to include and which parts, how many and which codification texts consult, which words to analyse, what counts as a word, what information to include from the codification texts, how to present that information and what to do with it. I consider all these issues under

their respective headings below: *Selecting a sample*, *Sources consulted* and *Coding and collating*. I end the section with a critique of the approach.

Selecting a sample

As I noted above, this approach to the data bears some relation to corpus linguistics and to variationist sociolinguistics, which are both concerned with representativeness and, consequently, sampling methods. Corpus linguists typically assume that their body of texts should be a standard reference for the language or language variety under investigation (McEnery & Wilson, 2001), devising strict sampling methods for ensuring this. Variationist sociolinguists typically focus on a hypothetical 'speech community', seeking to ensure that the body of texts to be investigated (e.g. sociolinguistic interviews) are representative of this wider speech community. However, despite the broad similarities, in some ways my priorities were different from those of corpus linguistics and variationist sociolinguistics. Since my analysis is quite labour-intensive, I have had to draw on quite a small amount of data. It is probably more accurate to say that my texts are representative of very specific types of discourse – the wider set of *Syrjeri* items and meetings at the housing association. They serve to illustrate the kinds of distribution of planned words that might come up in these situations. I do not claim that they are typical or generalisable to a wider population. Nonetheless, the institutional contexts serve as a wider frame for interpreting language practices.

As regards selecting a sample from the existing data, I decided to prioritise certain key characteristics rather than aiming to represent the wider dataset by randomly selecting samples, namely topic/type of vocabulary and participant type. I also prioritised selecting fairly long stretches of bounded text for the sample, rather than randomly selecting lines, or, indeed, picking out a specified number of words from a list of all the words that occur in the data (which would have made it difficult to know the denotational or referential meaning of the word-forms). Prioritising bounded text would also help me select talk that

revolved around particular topics, and avoid sections of general talk, such as greetings and introductions, and digressions (a salient feature in the workplace data, in particular). Whilst vocabulary in this kind of talk can also be viewed as part of the lexical planning project, broadly defined, non-specialist or non-context-specific vocabulary items are less likely to receive the specific attention of planners.

Deliberately sampling with an eye on participant type was motivated by a wish to incorporate an independent variable into the analysis. This required prioritising certain parts of the data over others, which definitively goes against the principle of representativeness. However, since I had not followed the principle of representativeness in other ways, I didn't find it particularly meaningful to insist on it in this case.

In the event, these considerations were borne out differently in the case of each dataset. Both sets are quite different in nature, in terms of the length of individual parts of the data and in terms of the focus of these parts. The media data consists of twenty focused items (all fairly repetitive in terms of their discourse structure), between 6 and 8 minutes long. This meant that I could select a certain number of texts where expert talk was prevalent and that all included all three participant types – presenters, expert and guests. I chose to select four texts, namely *Gallstones*, *Diabetes*, *Lupus* and *Meningitis*. This comes to a total of approximately 29 minutes. I chose this number of items since it was manageable whilst being long enough to offer plenty of data to work on.

One of my main priorities for the workplace data was to select a sample that was comparable in length to the media data (having started with the latter). The workplace data consists of four texts that, individually, are much longer than the individual texts in the media data. Therefore I had to select sections from individual texts, rather than full texts from the data. Another difference between the datasets was that the workplace data included a far greater number of

participants (34 compared with 17) which were much more difficult to categorise according to participant type. Moreover, whereas I had originally envisaged staff and tenants as comparable participant types, the sheer number of staff compared with tenants and the restriction of tenants to one recording made this impracticable. One comparison that I felt I could make, however, in order to draw in an independent variable, was to compare use across recordings. Talk across workplace texts was more likely to be differentiated than talk across media texts, because the composition of the texts (in terms of participants, topics, aims, structure, etc.) varied more. Consequently, I decided to take a sample from all four recordings from the workplace data, all of a similar length, selected for their inclusion of work-related talk. This gave me a sample of 30.5 minutes, with approximately 7.5 minutes taken from each recording.

Sources consulted

Selecting sources for consultation required considering which sources were most relevant. This involved, for example, considering which texts deal with the topics that came up in the data and were therefore most likely to include the lexis under consideration. Deciding which subject-specific texts I would consult followed a consideration of the words used in the data itself. These decisions, then, will be discussed in more detail in the individual analysis chapters, together with an outline of the fields covered in the talk and the vocabulary used. However, I also decided to consult some general codification texts – dictionaries and dictionary-like glossaries (that is, glossaries that include lexis on a number of subjects) – because of their applicability to a number of fields. Moreover, it is unknown whether speakers consult subject-specific codification texts (glossaries) or whether they consult general texts. I discuss the general texts here, all of which were consulted for both analyses. I also outline some general principles relevant to the selection of all codification texts.

The considerations in deciding which sources were relevant included: the likely authoritativeness of texts, their currency, the extent of their dissemination, and

their ease of use. On the likely authoritativeness of texts, since lexical planning is carried out by a number of agents there is no perceived central authority on lexical elaboration or codification more broadly, and there is even a degree of competition between planners (see Chapter 4). If so, which texts would count as most authoritative/legitimate? Or which would best represent the broader lexical planning endeavour? The answer partly lies in the other considerations I mentioned: currency, dissemination and ease of use. On the currency of texts, lexical planning has been carried out in Wales for many years. It goes back to the 16th century, but even if we were to only consider those texts that coincide with the deliberate efforts, since the 1950s, of extending the use of Welsh to new domains, such as education, there are some texts that are clearly more current than others. Are there some texts that are ‘too old’ to be considered? On dissemination, those texts where planners have made a discernible effort to disseminate their work (e.g. through marketing, tying in with authoritative organisations, or working with target audiences) are likely to have had more effect than others. Similarly, on ease of use, texts that are easily accessible are likely to have had more impact. In fact, when planners attend to accessibility and ease of use, it is with an eye on dissemination. In her model of an ideal term planning project, Bhreathnach (2012) argues that a number of factors contribute to the effective dissemination of terminology, one of which is an online terminology resource. “This means a simple, easy to use portal to a database which allows users to find the information they need quickly, and in suitable formats (direct search, downloaded lists, as resources for translation software, etc.)” (2012:101).

I have sought, then, to prioritise texts that can be considered to be authoritative, widely-known, current, available online, or at least in electronic format, and easy to use. I have also prioritised texts that are part of larger databases, since the integration of a number of codification texts in the same place contributes to their ease of use. Some of these factors tend to go hand-in-hand. For example, all electronic resources are fairly current, and their availability online, especially,

lends them greater visibility. This helps lend them authority (although not always). Nonetheless, it has not always been possible to satisfy all criteria at the same time, and at times I have prioritised some criteria over others.

Many of the texts I consulted have been compiled by the *Canolfan Safoni Termau/Centre for the Standardization of Welsh Terminology* (henceforth called the *Terminology Centre*) at Bangor University. The Terminology Centre is one of the most active players in Welsh lexical planning. It typically seeks to integrate some implementation measures into the design and dissemination of its terminology projects. It makes wide use of computer-assisted technology, which helps with the dissemination of terms as well as positions its work as modern and up-to-date. One of the ways it has used computer-assisted technology is through the integration of its glossaries in easily-accessed databases, most notably a searchable database that can be download to a computer desktop, *Cysgeir*. *Cysgeir* is part of the Welsh-language computer aid software, *Cysgliad*³⁷, which also includes the spelling and grammar checker, *Cysill*. *Cysgeir* houses a number of relevant glossaries and dictionaries, which I consulted for the analysis. I deal with these glossaries and dictionaries separately, because although *Cysgeir* collates the codification texts and can be considered a single resource, for each word it offers, it cites the source text. But when I do discuss the individual glossaries and dictionaries, it is worth bearing in mind that part of the reason I have consulted these particular texts is because they are housed within *Cysgeir*. I made extensive use of *Cysgeir* because the Terminology Centre is so active in lexical planning, because it is contemporary, and because it is so easy to use. The latter point relates both to the likelihood that the terms are fairly widely disseminated, and to my own convenience in conducting the analysis.

The *Cysgeir* programme houses the somewhat confusingly named *Cysgair* dictionary (Hicks et al., 1997) (note the small difference in spelling). *Cysgair* (the dictionary) is an English and Welsh dictionary, published in 1997 by the

³⁷ www.cysgliad.com

Terminology Centre, as a digital resource which had a very similar interface to the *Cysgeir* programme (it can be seen as *Cysgeir*'s precursor). The *Cysgair* dictionary is not described as a dictionary of terms, and it offers a number of synonyms in its results (which goes against the principles of monosemy and univocity that planners at the Terminology Centre typically seek to follow (see Chapter 2)). For example, a search for Welsh *rheoli* brings up English *control, curb, govern, manage, regulate, rule* and *run*. Nonetheless, I have included information from the *Cysgair* dictionary since it is so prevalent in the *Cysgeir* programme. It is unknown to what extent users of the *Cysgeir* programme will be aware of the differences between types of codification text (a point also made by Bhreathnach (2012: 101)), and it is probable that those who use the programme will take up the vocabulary items offered in the *Cysgair* dictionary.

One of the most comprehensive, easily accessible and apparently influential glossaries of modern Welsh terms is *Y Termiadur* (Prys, Jones, Davies, & Prys, 2006) (the word is a combination of *term* and *geiriadur* ('dictionary')). It is a glossary of terms compiled originally for use in education. It can be searched in English-Welsh or Welsh-English mode. Although it is a product of terminological work, its scope is general in the sense that it includes terms from a number of different subject-areas. It was compiled by the Terminology Centre for the *Qualifications, Curriculum and Assessment Authority for Wales*, and hence it is widely used in schools. I have included it for its extensive coverage and its influence, being used in education. It is also housed in the *Cysgeir* programme as well as in other electronic resources, including its own website³⁸.

Finally, *Geiriadur yr Academi* (Griffiths & Jones, 1995) (or *The Welsh Academy English-Welsh dictionary*, to give its English title) is an extensive English-Welsh dictionary, the most comprehensive of its kind for Welsh. Before the widespread availability of electronic resources it was probably the most widely used lexical codification text available for Welsh. It is still considered authoritative,

³⁸ www.termiaduraddysg.org

sometimes referred to as *Y Beibl* ('the Bible'). It is also commonly referred to as *geiriadur Bruce* (Bruce's dictionary) or simply *Bruce* (after the chief editor, Bruce Griffiths). According to Davies (2007: 673) it is "cyhoeddiad mwyaf ysblennydd y 1990au" ('the most magnificent publication of the 1990s'). Although one of its editors recently claimed that they had coined few neologisms for the dictionary (Jones, 2013), the lexicographical work is a contribution to the lexical elaboration of Welsh, as suggested in the preface of the dictionary, which lays out the need for Welsh to "offer a means of communication adequate to compete with English in every sphere of life" lest its speakers increasingly rely on borrowings and "end up speaking a patois with the feeling of linguistic inferiority which has given some Welshmen an excuse to abandon their native tongue" (J. E. C. Williams in Griffiths & Jones, 1995: vii). This quote could be read as an exposition of the thinking behind the whole lexical planning endeavour for Welsh.

Coding and collating

I discuss in the individual analysis chapters how I selected which words to compare against the codification texts, since the exact choice was quite specific to the data, and the discussion sits more comfortably within the analysis chapters. The ensuing discussion gives some information on how I dealt with the ambiguity of the 'word', and how I coded the individual items for consideration.

The words that I have selected for consideration are lexemes that are lexical in nature, what Carter (1987) calls *lexical items* (for Carter, lexemes can be either grammatical or lexical). By 'lexical', I mean that a lexeme carries propositional content (or the brunt of it). Lyons (1995) differentiates in this regard between 'full' (lexical) and 'empty' (grammatical) word-forms, arguing that full word-forms are more meaningful than empty, presumably where 'meaningful' relates to semantic/propositional meaning. A *lexeme* is the underlying conceptual sign that may be represented by a number of word-forms. For example, the English word-forms *types* (verb, 3rd person singular³⁹), *typed* (past participle) and *type*

³⁹ In 'Standard English'.

(3rd person singular) are all used in practice to represent the same lexeme (or lexical item⁴⁰), which can be labelled TYPE. This labelling method follows the standard practice of using the basic word-form⁴¹ in capital letters (Carter, 1987). However, I do not typically follow this method in analysis, unless to make explicit that I am referring to a lexical item.

I have, then, selected lexical items for inclusion and collated all the word-forms that represent each item, labelling them all with the lexical item label. All the word-forms include, for example, mutated forms, singular and plural forms (of nouns and verbs), verbs marked for tense, etc. To bring in an example from the data, the lexical item TENANT is represented in the Housing Association data by *tenant* (singular form), *denant* (mutated singular form), *tenantiaid* (plural form), *denantiaid* (mutated plural form) and *tenants* (other plural form).

However, deciding what words to code involves some other considerations, which are related to what is and isn't a lexical item. For example, as we see in the TENANT example above, there are some word-forms that are denotationally and grammatically equivalent, yet suggest different languages of origin. Does the use of both plural forms, *tenantiaid* and *tenants*, imply two different underlying lexical items (one Welsh and one English) that happen to look and sound the same, or does it imply competing strategies for plural marking for a single underlying lexical item? A similar problem arises with similar but differently pronounced word-forms. For example, /ɪn'ʃʊərəns/ and /ɪn'ʃiwrans/ might represent one or two lexical items, which we might represent as INSURANCE and INSIWRANS. I have dealt with this problem intuitively, deciding on a case by case basis whether I feel the word-forms are treated differently or would be perceived

⁴⁰ I will use 'lexical item' henceforth instead of 'lexeme', unless I want to include grammatical lexemes in my reference.

⁴¹ By 'basic' word-form, I mean word-forms that do not carry any grammatical markings. In Welsh, for example, these would include non-mutated forms, infinitive (in the case of verbs), singular, etc. In the case of some nouns, however, the plural, or collective, is the basic word-form, and the singular is formed by adding a singular suffix (e.g. *moron* ('carrots') > *moron-en* ('carrot')). In the case of some verbs, such as the loose equivalent of 'should', there isn't an infinitive (*dylwn i* ('I should'), *dylet ti* ('you should'), etc.).

differently by participants.

Another consideration was whether phrases consisting of multiple words should be considered single lexical items or not. Lexical items may be made up of one word or many. Even where a lexical item is made up of many words, it is lexically simple, in the sense used by Lyons (1995). That is, it is a vocabulary item, and is not constructed by the sentence grammar. Lyons explains with the example *red herring*: where it refers to ‘something that is pursued in vain’ the expression is lexically simple and therefore a lexical item. If, however, it means ‘a herring that is red’, it is lexically composite (formed by the sentence grammar), and therefore not a lexical item. I call lexical items that are made up of many words ‘multi-word items’ (following Moon (1997)). Moon uses three criteria to define multi-word items: they are institutionalised, that is there is some agreement within the ‘language community’ that they constitute a single lexical item; they are typically *fixed* as a sequence of words; and they are non-compositional, that is, they cannot be interpreted word-by-word. However, a number of words came up that were difficult to categorise according to these criteria. They seemed to be recognisable as fixed chunks, ‘fossilised’ as Moon calls it, and yet they could also be interpreted word-by-word. To take another example from the data, *bysedd traed* (‘toes’ in English, but literally ‘foot fingers’) is clearly interpretable word-for-word, at least if you think of toes as types of fingers, and yet is a fixed expression.

Moreover, there are instances in the workplace data of strings of words that are clearly fixed and very specific to the organisation (not the ‘speech community’, as Moon calls it, but the community of practice). These include, for example, job titles and department names, as well expressions used to refer to working methods or budgets, etc. Since these lexical items are so specific to the organisation, it is highly unlikely that they can be found in any codification text, and yet they do not necessarily go against the general principles underlying lexical planning. There isn’t an easy answer to this question. Wray (2002) concludes that there is no foolproof method for identifying formulaic expressions

(and we can consider multi-word items types of formulaic expression), all having their own particular difficulties. In the event, I relied on intuition and tended to be conservative.

After coding the relevant words in the data, I collated all codes (the lexical items) into a list to compare against the selected codification texts. Tables 3.1 and 3.2 show a section of the tables I drew up in consulting one codification text (Table 3.1 (media data) shows the table for *Geiriadur yr Academi* and Table 3.2 (workplace data) for *Termiadur*). Since I referred to a number of codification texts, the process was repeated for all texts with the sections for each text placed next to each other horizontally. The first column on the left shows the lexical items that are represented in the data through their word-forms⁴². Lexical items (those found in the data) may be Welsh or English (or ambivalent). The second column gives English translations of the lexical items, for the benefit of the reader. Sometimes this column includes a number of synonyms. ‘Translations’ given in this column for English or ambivalent lexical items are identical to the original, sometimes with minor spelling alterations (as in the case of AMBIWLANS). The third column, *E. entry*, lists the English headwords in the codification text where the lexical item can be found in the dictionary. In some cases, such as BUTTERFLY RASH, no English headword is listed (marked as *n*), because the codification text doesn’t deal with the lexical item. In such cases, all subsequent columns are marked as not applicable (*n/a*). The fourth column, *Listed W?*, notes whether the lexical item represented in the data is listed as Welsh in the codification text in question. If it is listed as Welsh, I have made a note in the next column (*W. spelling(s)*) of how it is spelled in that particular text, and in the next again (*Usage notes?*) of any usage notes. If the lexical item isn’t listed as Welsh in the codification text in question, I have marked *n/a* in both these columns.

⁴² Some columns are not shown here because of lack of space. I only show the columns that are needed for presenting the analysis. Columns excluded include, for example, a column loosely defining the lexical items, where this wasn’t clear.

Lexical item	E. translation (where applicable)	E. entry	Listed W?	W. spelling(s)	Usage notes?
AMBIWLANS	ambulance	ambulance	y	ambiwllans	n/a
ANYMWYBODOL	unconscious	unconscious	y	anymwybodol	unaware
AREN	kidney	kidney	y	aren	Anat.
ARTERY	artery, blood vessel, (poss. vein?)	artery	n	n/a	n/a
BUSTL	bile	gall, bile	y	bustl	bile > bustl (physiol.)
BUTTERFLY RASH	butterfly rash	n	n/a	n/a	n/a
BYS	finger	finger	y	bys	n/a

Table 3.1: Example of information collated from *Geiriadur yr Academi* on the media dataset

Lexical item	E. translation (where applicable)	E. entry	Listed W?	W. spelling(s)	Usage notes?
SYSTEM	system	system	y	system	(of body organs, computers, numbers)
TAFLEN	leaflet	leaflet, handout	y	taflen	leaflet (= of paper)
TÂL	payment	payment	y	tâl	(yn gyffredinol)
TALU	pay (v)	pay	y	talau	n/a
TARGED	target	target	y	targed	n/a
TENANT	tenant	tenant	y	tenant	n/a

Table 3.2: Example of information collated from *Terwiadur* on the workplace dataset

From these tables, the main information I have used is whether or not the lexical item is listed as Welsh. However, I use the additional information to contextualise in determining whether the lexical item can be considered codified. Not all items listed as Welsh are presented equally in codification texts, nor do they all relate equally to the data. For example, there are cases where it is unclear whether the pronunciation found in the data is represented by the orthography used in the codification text. This might occur where the spelling of a word is made to satisfy the conventions of Welsh orthography, but where pronunciation in the data has not been integrated. Or it happens where the spelling represents more than one possible pronunciation. An example from the data that combines both issues is

MENINGOCOCCUS, spelled in *Geiriadur yr Academi* ‘meningococws’. It is unclear whether the <g> represents, in the minds of the dictionary’s authors, /g/ or /dʒ/ (or possibly both), but the final vowel has been integrated, to suggest the pronunciation /ʊ/. In the data it is pronounced with /dʒ/ and /ə/ respectively (a not particularly Welsh pronunciation). In other cases, the lexical item represented in the dictionary includes additional or fewer grammatical words than are used in the data. For example, TRAWIAD Y GALON (‘heart attack’) is given in all codification texts as *trawiad ar y galon*. There are also cases of semantic ambiguity, such as in the case of ARTERY, which is sometimes used in the data to refer to ‘blood vessel of any type’, although in technical use it typically refers to a specific type of blood vessel. It is unclear from the codification texts whether they allow *arteri* as a superordinate term. Sometimes the usage notes shed light on meaning but, more often than not, the user has to rely on his/her intuition. In yet other cases, the lexical item is used by participants in the data with a meaning clearly different from that stipulated in the codification text, such as BUSTL, which is used by participants to mean ‘gallbladder’, but which codification texts list as meaning ‘gall’ or ‘bile’. I have marked most of such cases as ambiguous (or ‘partly codified’).

So far, then, I have three categories of codification – codified, not codified and partly codified (I summarise below). A fourth category is ‘not applicable’, where the lexical item is not dealt with at all by the codification text. This is the case with CHECIO (‘to check’), for example, where it is not codified as Welsh by any of the codification texts consulted, and so can be put in the ‘not codified’ category. However, there are some cases where there is inconsistency between codification texts. For example, *Geiriadur yr Academi* accepts *insiwrans* (‘insurance’) as Welsh. However, the other codification texts don’t accept it as Welsh at all. I have assigned such examples to a ‘mixed’ category. Finally, where a lexical item is assigned to ‘not applicable’ for one or more of the texts, but not for other texts, the ‘not applicable’ category becomes invisible, so to speak. For example, CLERC GWAITH (‘works clerk’) is only applicable to *Geiriadur yr*

Academi (where *goruchwyliwr gwaith*) is preferred. The concept is not dealt with in any of the other codification texts. These texts become irrelevant in the collated category, and CLERC GWAITH becomes ‘not codified’, solely on the basis of the *Geiriadur yr Academi* entry.

To recap the codification categories and summarise what they mean, with the proviso that ‘not applicable’ becomes invisible:

- i. codified: all texts show the lexical item as Welsh without ambiguity;
- ii. not codified: none of the codification texts list the lexical item as Welsh, preferring an alternative;
- iii. partly codified: the codification texts list a similar word-form as Welsh, which may or may not represent the lexical item found in the data;
- iv. mixed: there is inconsistency between codification texts;
- v. not applicable: none of the codification texts deals with the concept.

The distribution of these categories, then, provide a picture of the data, which seeks to draw out the extent to which participants use language that is codified in selected dictionaries and glossaries. Further, I sought to break down the data to see whether it offered a view on contextual or speaker variation, comparing the distribution of the codification categories within extralinguistic factors. How I sought to do this – the extralinguistic factors I considered – was dependent on the datasets themselves. Being small datasets and not selected specifically with an eye on an analysis of extralinguistic factors, they afforded only some, data-specific, opportunities for this kind of analysis. Since my choice of extralinguistic factors required a contextualised understanding of the datasets themselves, I discuss my choice of these factors in the individual analysis chapters.

Critique

Following variationist assumptions, this part of the analysis seeks to explain variable linguistic practice by highlighting differentiated patterns of distribution,

characterised by certain extralinguistic factors. Although this approach to language in society has been fundamental in the development of sociolinguistics, it has been critiqued. Cameron (2009 [1990]) underlines that finding a differentiated pattern of distribution is not the same as explaining that difference. She notes that sociolinguists have typically invoked ad hoc sociological theories to explain distributional patterns. A pattern of distribution cannot be explained simply by invoking a concept such as identity. It is not enough to say, for example, that working-class teenage women from south-west England use rhotic /r/ a certain percentage of the time because they are playing out their identities as working-class teenage women from south-west England. This would be a circular explanation, which Cameron calls the *correlational fallacy*. It doesn't explain why. Similarly, if we find a difference between, for example, the lexical use of presenters and guests in the media data, what does this mean? This is not to say that such studies have no value – at the very least they can highlight where 'something is going on' in the data, but they cannot satisfactorily explain what that something is without a more robust theoretical approach. We are in the realm of social meaning again.

I will explain patterns of distribution partly through a more thorough understanding of the sociolinguistic and institutional context. Cameron (2009 [1990]) argues that conventions for the use of language do not evolve organically (the *organic fallacy*), but are produced by people and institutions, as we found in the discussion on standardisation in the previous chapter. Languages, she insists, are regulated social institutions (2009 [1990]:113). Moreover, linguistic resources aren't freely available to speakers as closed systems, but are distributed unevenly. This is clear in the case of lexical planning in Wales – only to those who access planned language is it available for deployment in interaction⁴³. Considering the sociolinguistic and institutional framework, then, a task that I undertake in the next chapter, will help shed light on what lexical

⁴³ Lest I paint a picture of those who do not have access to 'standard' language as deprived, however, it ought to be pointed out that they will have access to other linguistic resources that 'standard' speakers may not.

resources are available and what meanings are available in their deployment.

Beyond the problem of finding compelling explanations for distributional patterns, however, variationist sociolinguistics has been critiqued from a different angle. Statistical methods simply cannot account for all of variation and its meaning. Quantitative methods search for trends, and run the risk of dismissing as ‘noise’ data that doesn’t come within 95% or 99% of the normal curve. Coupland (2007: 41) argues that even single occurrences of a single variant can be socially meaningful, the 5% or 1% at the margins of the normal curve, we might say. The problem, he says, is in the quantitative design. Whilst valuable in some ways, variationist designs are too restrictive because of their need to shape linguistic and social data into linear strings (2007: 41), whereas language isn’t inherently linear.

Finally, in my discussion on the implementation of language planning in Chapter 2, I argued that it is at the level of practice that implementation occurs. I also suggested that implementation isn’t necessarily a clear case of language planning being implemented or not, or of acceptance of the planned language or its rejection. Equally possible is for various ‘policies’ – informal and covert – to be in conflict in situ. A quantitative account won’t give us an impression of conflicting policy, since it tends to lump together language practices in a normative fashion, which glosses over potentially meaningful differences. We get a generalised account of language practice, without getting a view of the informal/covert, potentially conflicting, language policies that are operational in talk. This is precisely because quantitative accounts are typically descriptive, rather than explanatory. The next section introduces the interactional analysis, which seeks to view the data from a different perspective, in order to address these issues. Although both parts of the analysis are contrastive, in that they take a different approach to sociolinguistic variation and how it can be studied and explained, it is hoped that together they are complementary.

3.6 IN SITU LANGUAGE POLICY/PRACTICE

Thinking about language policy variously as formal/informal and overt/covert (see Chapter 2) is a useful way of reconceptualising the work of language planning implementation and of emphasising the fundamental role of in situ communication and language practice. On the other hand, at the level of interactional practice, informal and covert ‘policy’ relate quite directly to the subject matter of other broadly sociolinguistic research approaches, notably conversation analysis (CA), interactional sociolinguistics (IS), critical discourse analysis (CDA) (including the study of ideologies) and the analysis of sociolinguistic style (Coupland, 2001, 2007; Kiesling, 2004). All these approaches, being closely related historically, share a constructionist perspective on language and the common goal of explaining why speakers use the linguistic forms they do in situ⁴⁴. All of them see communication as being strategic in some sense. Gumperz (1999: 454) writes from an interactional sociolinguistic perspective:

“Communicative practice largely rests on the discursive practices of actors acting in pursuit of their everyday goals and aspirations. Speaking, when seen in practice perspective, is not just a matter of individuals’ encoding and decoding of messages. To interact, as conversational analysts have shown, is to engage in an ongoing process of negotiation, both to infer what others intend to convey and to monitor how one’s own contributions are received. In other words, at issue are shared interpretations rather than just denotational meaning.”

I think all the approaches I mentioned would subscribe to some extent to this view, although they do differ in terms of the degree of rational, conscious choice

⁴⁴ Stubbe et al. (2003) is a useful comparative study of broadly interactional approaches, that analyses a single piece of interaction from five different perspectives: conversation analysis, interactional sociolinguistics, politeness theory, critical discourse analysis and discursive psychology.

they allow for on the part of speakers. For example, from his critical perspective, Fairclough argues that language practices are driven by ideologies, and that, since ideologies are typically below-the-radar (*covert*), “self-conscious linguistic choice is a relatively marginal aspect of the social processes of text production and interpretation” (Fairclough, 1995b: 18). From a critical perspective, then, it is ideologies that are played out in communicative practice, and not the individual’s ‘everyday goals and aspirations’. It should also be emphasised, however, that even in research approaches that focus more on the individual’s goals, intentionality isn’t always viewed as conscious or rationalised. This is a point made fairly early, in Blom & Gumperz’s (1972) study of code-switching (or style-shifting) in Norway, where they found that, on playing recordings back to participants who had code-switched on tape, those same participants were quite surprised to find that they had switched, and even vouched to stop doing so in future. When I say that these approaches view communication as strategic, then, I do not necessarily mean that action is always teleological. The main point I want to make in relation to strategy, however, is that strategy relates directly to policy, and that consequently all these research approaches relate, in some way, to what LPP researchers might call informal, sometimes covert, policy.

In their explanations, all these approaches seem to take as given that there is linguistic and communicative variation, which supplies speakers with options for making meaning, or with a sociolinguistic repertoire. As Fairclough argues, selecting from these options constitutes more than merely choosing one linguistic form over another (say, a native word over a borrowing). Rather, “these formal choices constitute choices of meaning, the selection of options from within the *meaning potential* – how to represent a particular event or state of affairs, how to relate to whoever the text is directed at, what identities to project” (Fairclough, 1995b: 18, my emphasis). Meaning potential, however, is a concept used to emphasise that the relationship between sign and social meaning is fluid. There is an assumption that there are typically normative meanings within cultural configurations, and that these cultural configurations typically provide a

socio-cultural frame for understanding, or inference⁴⁵. However, 1) talk can occur between people who have divergent socio-cultural frames (a specific focus of IS); 2) within socio-cultural frames there are other frames, such as genre, which can also provide their own resources for meaning-making; and 3) even from within a single socio-cultural frame, normative meanings can be contested.

“We will need to think in terms of social *meaning potential* [...] being called up or activated or validated, or undermined or challenged or parodied, in particular discursive frames for particular local effects. This would imply, once again, that social meaning doesn’t exclusively reside in linguistic forms, or even in so-called speech communities or in speakers’ sociolinguistic histories and experiences. It is partly a situated achievement in acts of speaking.”
(Coupland, 2007: 24)

Where the approaches I mentioned do differ, however, is how they explain how and why speakers make these choices, that is, in terms of what counts as a valid interpretation. Conversation analysts, for example, (claim to) ensure the empirical groundedness of their interpretations by avoiding claims that cannot be supported by explicit evidence found in the data. In contrast, interactional sociolinguistics “allows us to take explicit account of the unstated assumptions and background knowledge the participants in an interaction bring to bear as part of the interpretive process” (Stubbe et al., 2003: 362). Interpretations of participants’ unstated assumptions and background knowledge can be very hard to support empirically. In practice, however, conversation analysts arguably also make inferences about background knowledge in their interpretations of conversation participants’ communicative behaviour, but this isn’t built formally into research design. For example, in his conversation analysis of an instance of code-switching in an interaction between English/Cantonese speakers in Tyneside (England), Li Wei (2002: 168) infers that the reason a female

⁴⁵ The influence of the socio-cultural isn’t emphasised as much in conversation analytic research.

participant switched to Cantonese when addressed in English by a younger male was that “she belonged to a generation whose language choice and language preference [...] were both clearly Chinese-dominant”. Li Wei’s analysis does acknowledge “attitude, [language] preference, and community norms” (2002: 168), since it seeks to show how these are ‘brought about’ in interaction. However, there is no systematic approach to knowing what those ‘attitudes, preferences and community norms’ are, other than what is possibly evidenced in the interaction. I say ‘possibly’ because it is by no means clear that language preference was the reason for the switch. It is an inference made by the researcher based on episode-external knowledge.

Perhaps it is a matter of what researchers find interesting and research-worthy. Li Wei claims that his interpretation “told us little that we did not *intuitively* know already about bilingual speakers’ language behavior” (2002: 168, my emphasis), thus relegating ‘attitudes, preferences and norms’ to the realm of the obvious and non-researchable. It ought to be acknowledged, however, that it is Li Wei’s prior knowledge of the Tyneside Cantonese/English community that gives him this ‘intuition’, gained through years of research, including a three-year period of residence in Tyneside, during which he experienced first-hand Tyneside Chinese adults’ refusal to speak to him in English (Li Wei, 2000). Knowledge of this sort is not universal. In interactional sociolinguistics and critical discourse analysis this kind of knowledge is formalised within the research design. ‘Intuition’ becomes ethnographic-like knowledge of the background to the speech event or Fairclough’s (1995b) ‘socio-cultural goings on’, or the detailed accounts of social, economic, political and geographic conditions found in anthropological linguistics (such as Hill, 1985). It is not that CA cannot be of use. Indeed, CA can show how ‘attitudes, preferences and norms’ are brought about in interaction, and key CA concepts, such as turn taking, are used in other research approaches. But I agree with the perspective taken by interactional sociolinguists and critical discourse analysts that we need to pay attention to how we, as analysts, infer what participants’ attitudes, preferences,

norms, background knowledge and unstated assumptions are. There is inevitably a degree of indeterminacy in making interpretations about such concepts. But this doesn't mean that we cannot analyse from such a perspective. Rather, what is primary is that, as researchers, we are reflective and critical of our own interpretations.

To summarise then, in considering participants' in situ language 'policies', I turn to various related interactional approaches to language as a communicative tool, specifically looking at participants' strategic use of language in achieving their goals, but allowing that such strategies might not be deployed consciously or rationally, and that they may be ideologically founded (in fact, I do use ideology as an analytic tool – see below). 'Strategies' involve participants selecting from various options in their linguistic repertoire and deploying them to activate one, or perhaps more, of their potential meanings. Finally, inferring what participants mean necessitates making inferences about participants' background knowledge, attitudes, norms and ideologies. I seek to make these inferences from considering the conversational structure, but also by paying attention to talk in the wider dataset (not from the particular sequence in question), as well as by considering the institutional and sociolinguistic context (this I detail in Chapter 4).

Social meaning and bilingualism

Whilst my research focuses primarily on language planning and standardisation, I also need to consider social meaning in relation to bilingualism. As I argued in Chapter 2, lexical planning is presented as a solution to widespread borrowing. It can be viewed, then, as a primarily purist effort. I have argued elsewhere (Robert, 2011) that purism appears to be central to Welsh standard language ideology, and purism is an ideology that responds to – and seeks to restrict – the effects of language contact⁴⁶. I have not wished to focus my analysis entirely on

⁴⁶ Not all researchers restrict purism to contact situations. For example, Wexler writes that "The object of elimination is often foreign elements which purists seek to replace by native elements, but there can also be purism directed against elements coming from within, such as geographical and social dialectalisms, archaic elements or neologisms" (Wexler, 1974: 1). In my work,

bilingualism and the code-switching literature, because there are issues that pertain to language planning and standardisation in Wales that are not directly or only related to bilingualism. For example, Lloyd-Morgan (2010) considers a perhaps unlikely conundrum: “What’s Welsh for *woman*?”, presenting a number of possibilities, all native Welsh words (e.g. *dynes*, *merch*, *menyw*, etc.), but concluding that none works as a cover-all pan-Wales word that avoids what she calls ‘unwanted connotations’ (unwanted from her feminist perspective). This is a matter that pertains directly to standardisation, but not directly to purism or bilingualism. However, in the analysis itself bilingualism emerges as a prevalent theme. We will see, for example, that there are very few subject-specific words in the datasets that can be said to be regionally-specific – that have the potential to index specifically regional identities. But there are a number of words that relate in some way to language contact, having the potential to index social identities within a purist interpretive framework. For example, there are a number of cases of what I call *lexical variation* in the data (where two or more lexical designations are used within a dataset as (broadly) equivalent semantically), the vast majority of which involve at least one word that can be said to be contact-induced. Bilingualism, then, deserves some consideration. In this section I will briefly outline how I approach social meaning specifically in relation to bilingualism and language contact phenomena.

Considering my focus on *lexical* planning, it is lexical borrowing specifically that is relevant to my analysis (rather than other kinds of contact phenomena). Researchers disagree as to whether lexical borrowing is distinct from code-switching. I tend to agree with Poplack et al. (1988), who reserve borrowing to refer to *lexical* contact phenomena and use code-switching to refer to sentence-level contact phenomena⁴⁷. Poplack and her colleagues postulate that the status of borrowings can vary – there are established borrowings (or *loanwords*) and

however, *purism* refers to an ideology that seeks to restrict elements that are *perceived to be* foreign.

⁴⁷ This definition is somewhat complicated by the fact that some sentences consist of a single word. This seems to be what Poplack et al. are referring to when they postulate that there are single-word switches, but they do not elaborate or give examples.

infrequent or first-time (*nonce*) borrowings⁴⁸. Loanwords are defined as borrowings that “recur relatively frequently, are widely used in the speech community, and have achieved a certain level of recognition or acceptance, if not normative approval” (Poplack et al., 1988: 52). As we see in this definition, loanwords aren’t necessarily universally established. What might be called a loanword on the basis of frequency, for example, may or may not gain ‘normative approval’. The borrowing *meddyg*, for instance, is accepted as Welsh by lexical codification texts. However, its semantic equivalent, *doctor*, although by no means a ‘nonce’ borrowing (it is arguably more widespread than *meddyg* and it dates back to at least the 14th Century (*GPC*)), is accepted by only some.

For some researchers (Gardner-Chloros, 2009; Myers-Scotton, 1993) non-established borrowings are best viewed as a type of code-switch, on the assumption that until a word is well-established, by some measure or another, the speaker must be switching into the other language in order to access it. For example, Gardner-Chloros (2009: 31) assumes that “loans must start life as code-switches and then generalize themselves among speakers of the borrowing language”⁴⁹. This view seems to me to be premised on the assumption that a bilingual has two distinct lexicons serving two distinct linguistic systems. Other code-switching researchers (Auer, 2007; Heller, 2007; Woolard, 1998) have questioned the validity of this assumption arguing that much of bilingual practice

⁴⁸ For Poplack and her colleagues, ‘nonce’ and ‘established’ are, in fact, operational categories – borrowings that occur only once in their corpus, and borrowings that are used by many speakers, respectively. They seem to infer that this relates to differences in the status of borrowings (established or not) for speakers, although they do include the proviso that “*a priori* we cannot take for granted that this purely operational distinction correlates either with the degree of linguistic assimilation of the word or with its history of attestation as a loanword in the language” (50). Nonetheless, they do seem to infer that there is a difference between words that are borrowed only once or for the first time and borrowings that are used frequently, culminating in their acceptance in a lexical codification text.

⁴⁹ In fact, this view seems to be very similar to that of Poplack and her colleagues that established borrowings start out as nonce borrowings. The difference between Poplack and Gardner-Chloros, it seems to me, is how they categorise lexical contact phenomena. For Poplack, all lexical contact phenomena, established or not, are borrowings, code-switching being restricted to sentence-level contact phenomena. For Gardner-Chloros, on the other hand, established lexical contact phenomena are borrowings, but non-established ones are switches. There doesn’t seem to be a major difference between their views that well-established lexical contact phenomena, whatever label they are given, were once transient borrowings and became established through increased use.

involves syncretic or bivalent language use. In agreement, I prefer to think in terms of a bilingual possessing a single linguistic repertoire, with linguistic resources with different potentials to signify, and differently activated according to need or context. As Heller (2007: 8) asks rhetorically:

“what if we replaced the idea of code with the idea of linguistic resources which are socially distributed, organized certainly by speakers individually and collectively, but which do not necessarily ever have to correspond to some closed and wholly describable system? What if language were part of a set of practices which had varying manifestations (both for individuals and sets of networks of people), but which could not be firmly distinguished from other kinds of behaviour? What if grammar were the order speakers impose, more or less successfully, on their linguistic resources?”

From this perspective the question raises itself whether it matters whether a word or linguistic construction is a switch or a borrowing by some formal definition. The point, I think, for my research is whether contact-induced lexical items are viewed by speakers (consciously or not) as borrowings, as appropriate to the situation or to the speaker's needs, as well-established, as ‘Welsh’, as filling ‘gaps’, etc. “What is the differential status (if any) of objects such as ‘one language’ and ‘the other’ in speech behaviour in relation to activities and identities? To what extent do materials in ‘language A’ or ‘language B’ *count conversationally as* materials in ‘language A’ or ‘language B’, and not as something else?” (Alvarez-Cáccamo, 1998: 41). That is, what do borrowings mean socially in interaction?

In asking these questions, all borrowings are of interest, since within a purist ideology, which is central to Welsh standard language ideology, all borrowings, be they well-established or ‘nonce’ or ‘switches’, however we as analysts choose to define them, are potentially fair game as targets for modification or avoidance.

Certainly we can argue that not all borrowings *are* targeted within a purist ideology all the time, but they are all open to it. In his descriptive framework of lexical purism, G. Thomas (1991: 73) differentiates between lexical items that have been ‘assimilated’ (which I assume to be synonymous with ‘integrated’) and those that have not. That is, Thomas differentiates between borrowings that have more or less potential to mean ‘own-ness’ and ‘otherness’. For example, Thomas differentiates between ‘mild’ purism, such as opposition to a recent English borrowing in German like *Marketing*, and more ‘extreme’ purism, such as opposition to well-established and integrated borrowings like German *Zigarre* (‘cigar’) (1991: 172). The former may be defined in analytic terms as a ‘nonce’ borrowing or a code-switch and the latter as a well-established borrowing or a ‘loanword’, but they are both, nonetheless, interpretable as borrowings, and within a particular ideological orientation, possible targets of purification.

Thomas’ view, nonetheless, is similar to Poplack’s and Gardner-Chloros’ that there is some potential difference between lexical contact phenomena, depending on the degree of integration or some other measure. I do not altogether disagree with this view. Bilingual resources, including lexical variants (different words for making the same semantic reference, such as *meddyg* and *doctor*) and methods of integration, provide potentially meaningful resources for bilinguals to deploy in interaction. (I shall outline in Chapter 4 what kinds of linguistic resources are potentially meaningful in the case of Welsh/English bilinguals). But I am sceptical as to how fixed meanings are. Specifically, I am sceptical that there can be a fixed degree of establishment. If all borrowings are open to interpretation as borrowings, then whether they are indeed interpreted as borrowings depends, to some extent, on ideological perspective. Where ideologies change (over time or across sociolinguistic contexts), the meanings activated by borrowings change too.

Perhaps a better notion than ‘well-established’ is that of normalisation, a notion that we have come across already in my discussion on standardisation. I will turn

to normalisation next, not only in the light of bilingual resources for meaning-making, but in the light of standardisation more generally. As Woolard (2005) has argued in relation to code-switching, from the perspective of a shared linguistic repertoire, it is just as pertinent to ask why a speaker does *not* switch as it is to ask why they *do*. This is precisely the focus of my research, of course, in as far as bilingualism and purism are concerned: when speakers use ‘Welsh’ words (but defined as Welsh by what criteria?) why do they do so, and if they use borrowings, why? In order to consider social meaning in relation to standardisation and bilingualism, I need an approach to analysing normalisation in interaction.

Social meaning and normalisation

As we saw in Chapter 2, planned standardisation seeks to normalise the targeted language, so that it can be used by any two speakers without the assistance of and outside the constraints of the situation, to paraphrase Bourdieu (cited in Chapter 2). When I ask, then, about the social meaning achieved by words used in interaction, I am primarily concerned with whether they are presented as normal. Where normalisation is sought, the social and economic conditioning of language acquisition and use are played down or obscured, an ideological process akin to Gal & Irvine’s ideological semiotic process of *erasure* – “the process in which ideology [...] renders some persons or activities or sociolinguistic phenomena invisible” (1995: 974). Another concept that is closely related to normalisation is *naturalisation* (Fairclough, 1995a). For Fairclough, naturalisation pertains to ideologies, or more accurately, to ideological-discursive formations (IDFs), a concept he uses to draw attention to his view that ‘ways of seeing’ (ideologies) and ‘ways of talking’ (language practices) are inseparable (1995a: 40). The more naturalised an ideology, the more taken for granted and opaque it is. That is, the more the language practices by which it is expressed are taken to be commonsensical – not ideological, but an objective reflection of reality. The concept of naturalisation relates well to standardisation and normalisation, since I view standardisation as relating to an ideology that

obscures the potential for linguistic resources to mean in any way other than in a purely denotational sense (and where denotational meaning is taken to be easily defined objectively). The social meaning of forms that are considered ‘standard’ is rendered invisible – their ideological basis becomes opaque.

Naturalisation is also a useful concept for my research since Fairclough theorises naturalisation within *institutional* frames, which relates directly to my research perspective. His reason, he says, is that his broader critical research aims can only be approached within a framework that integrates micro and macro perspectives (interaction and language practices, and the sociocultural, respectively), and that institutions offer the best means of achieving this integration. Fairclough (1995a: 38) takes a broad view of institution, not merely as ‘organisation’, as the word is typically understood in lay terms, but as:

“a sort of ‘speech community’, with its own [...] set of speech events, its own differentiated settings and scenes, its cast of participants, and its own norms for their combination – for which members of the cast may participate in which speech events, playing which parts, in which settings, in the pursuit of which topics or goals, for which institutionally recognized purposes.”

An institution, then, can be a formally established ‘organisation’, with explicitly stated rules of engagement and structure, even a legally recognised entity within the wider institution of the state. Or an institution can be an informal collection of people and practices, such as a family. It is the institution, Fairclough argues, that enables and constrains the expression of ideology through discourse: “it provides [its members] with a frame for action, without which they could not act, but it thereby constrains them to act within that frame” (1995a: 38), to the extent that the frame offers clear norms according to which action is constrained. Sometimes there are competing norms within an institution.

The question of competing norms relates to another key concept of Fairclough's that is connected to naturalisation, namely orderliness. Orderliness is a useful analytic concept, since it can be operationalised to investigate naturalisation, and normalisation, in spoken interaction. According to Fairclough (1995a), within any institution there are multiple ideological-discursive formations in competition, but typically one will be dominant. Where an ideology is dominant, and consequently maximally naturalised, Fairclough posits that interaction will be orderly. The orderliness of interaction relates to "the feeling of participants in [the interaction] [...] that things are as they should be, i.e. as one would normally expect them to be" (1995a: 28). He stresses, however, that orderliness pertains to the participants' feeling of the interaction, and not to the analyst's. In considering, then, whether words in interaction are presented as normal, we can ask whether they are used in interaction that is orderly – that the participants feel to be 'as it should be'.

Participants infer (not necessarily consciously) whether interaction is orderly on the basis of their background knowledge. For Fairclough (1995a), background knowledge subsumes ideological representations (he is critical of what he calls 'descriptive' discourse analysts that they take background knowledge to be knowledge that is directly accessible to participants, and consequently neglect the ideological, which is typically opaque). He posits four main dimensions of the 'knowledge base' (and all four include ideological elements): knowledge of principles and norms of language use, of situation, of the world, and of *language codes*. On language codes, and lexis specifically, he argues that lexicalisations always have potential alternatives (different ways of being formulated), hence express different ideologies (on the assumption that 'ways of speaking' and 'ways of seeing' are inseparable).

"Alternative lexicalizations are generated from divergent ideological positions. And lexicalizations [...] may be more or less naturalized: a lexicalization becomes naturalized to the extent that 'its' IDF

achieves dominance, and hence the capacity to win acceptance for it as ‘the lexicon’, the neutral code.” (1995a: 34)

In Fairclough’s examples, alternative lexicalisations ascribe alternative identities primarily to the referent (he offers some alternative lexicalisations relating to youth, such as *irrepressible/incorrigible*, *debunking/defiance*). But alternative lexicalisations can also relate to the identity of the speaker (using ‘correct’, ‘pure’ ‘native’ words or using ‘messy’ ‘incorrect’ ‘borrowings’) or to the speaker-recipient as a collective, and thereby to the interaction order and more broadly to the social order.

Although Fairclough insists on taking the speaker’s view of what is ‘appropriate’, as analysts we are left with the task of inferring the participants’ knowledge base and their feeling for the orderliness of interaction. Fairclough undertakes this task partly by drawing on discourse- and conversation analysis – considering the participants’ ‘interactive behaviour’. Orderliness might relate, for example, to coherence of interaction, where speaker turns fit ‘meaningfully’ together, or to talking in an ‘appropriate’ way (which includes the use of ‘appropriate’ lexicon). Given that ‘appropriateness’ is to be seen from the perspective of the participant, however, the analyst must draw on some knowledge base other than the conversational structure itself (see my discussion on the validity of analysis above). If the use of lexicon in interaction is counter-normative, against what norms is the analyst comparing this counter-normative practice? There is, then, some reliance on knowledge that is garnered from outside of the particular speech event or interactional sequence under scrutiny, a knowledge that can be gained by turning to other speech events within the same institutional discourse (the wider dataset) (de Rooij, 1998) and/or by consulting other sources of knowledge regarding the institution. In my case, especially since my focus is on LPP within institutional settings, other sources of knowledge include formal policy documents. I also take any observations I made in my dealings with the organisations in which my research is located as another source of knowledge, as

well as other academic research on these or similar institutions in a Welsh-language context. There is not necessarily a direct relationship between what is inferred from these other sources and the ideological-discursive formations that are reproduced in institutional interaction. For example, the explicitly stated formal language policy found in S4C's language guidelines (S4C, 2008) is not necessarily a direct reflection of the 'covert, informal policies' followed in the broadcast data. But these other sources of knowledge can be taken as aids to the interpretation of orderliness and social meaning in the primary data. It is to these other sources of knowledge that I turn in my next chapter, considering the wider Welsh sociolinguistic context and the institutional context of each dataset.

CHAPTER FOUR:

The Welsh sociolinguistic context

4.1 INTRODUCTION

In this chapter, I present some detail of the sociolinguistic situation specific to Wales and the Welsh language that are relevant to my analysis of language practices. I start by giving some details on historical language planning and standardisation. Historical events are relevant, since they influence present-day ideologies and patterns of language acquisition and practice. Looking at historical planning opens a window on the kinds of lexical resources that are available and what meanings they may carry today. Secondly, I give some detail on contemporary sociolinguistic variation, in particular as regards lexis, before moving on to outlining contemporary attempts at lexical planning in more detail, and the governmental framework. Finally, I consider the institutional frames of the two datasets, to give an impression of the kinds of norms that we might see in operation.

4.2 A BRIEF HISTORY OF WELSH STANDARDISATION AND LEXICAL PLANNING

The planning and standardisation of Welsh are not new developments. Welsh has been undergoing a process of standardisation for centuries, at least as far as the written language is concerned and allowing that this process has not been a straight road⁵⁰. This planning and standardisation has involved a fair amount of lexical planning. In her historical study of Welsh lexicography, M. Morgan (2002) lists 38 dictionaries compiled (not all published) between 1547 and 1914. This figure does not include Biblical dictionaries, dictionaries of terms, dialect dictionaries, or vocabularies for tourists. Lexical planning, then, has a long tradition in Wales, with the first printed Welsh-English dictionary, William

⁵⁰ Differing language-ideological viewpoints resulted in different opinions as to how Welsh ought to be standardised and consequently some changes in direction, especially as regards spelling. See D. J. Jones (1988), for example, on language planning reforms driven largely by John Morris-Jones at the beginning of the 20th Century.

Salesbury's *A Dictionary in Englyshe and Welsh...*⁵¹, being published in 1547. Salesbury's dictionary was published not long before Bishop Morgan's 1588 translation of the Bible, an event that is largely held to have effected the standardisation of Welsh (C. H. Thomas, 1982 holds this view, for example). Welsh lexicography started in earnest, then, around the same time that other processes of standardisation took hold, spurred, no doubt, by the development of the printing press⁵².

Although the translation of the Bible was ordered by the English queen, Elizabeth I, as a pragmatic step towards unifying her realm (she was not concerned with the Welsh language per se), early lexicographers seem to have come from a somewhat different ideological perspective. As one of the leading Renaissance figures in Wales, Salesbury was keen to increase the standing of Welsh as a language of learning and 'high culture' (M. Morgan, 2002), that is, of elite knowledge. This spirit seems to have persisted in the years that followed, as we see from the title of Thomas Jones' 1688 dictionary, *Y Gymraeg yn ei Disgleirdeb / The British Language in it Lustre* (or 'brilliance') (T. Jones, 1688). Jones considered that current Welsh practices lacked dignity and refinement, and sought to "trosglwyddo gwybodaeth ieithyddol i'r di-ddysg a sicrhau eu bod yn trysori'u mamiaith" (M. Morgan, 2002: 47)⁵³. There seemed to be an assumption here that, in order to thrive, Welsh needed to have a specific form and had to be used, or usable, for all human pursuits, in particular for learned functions. The language of ordinary people (the 'uneducated') was considered to be unworthy/unable to fulfil such functions. Moreover, as we see in the title of Thomas' dictionary, the language presented was positioned as part of a great Welsh tradition that ostensibly dated back (as we see in the English translation of *Cymraeg as British*) to pre-Anglo Saxon times: Welsh as the authentic native

⁵¹ *A Dictionary in Englyshe and Welsh moche necessary to all suche Welshmen, as wil spedly learne the englyshe to'gue* (Salesbury, 1547)

⁵² In fact, Salesbury was both lexicographer and translator of religious texts, including the New Testament.

⁵³ 'pass linguistic knowledge on to the uneducated and to ensure that they treasure their mother tongue.'

British language.

With the significant social and epistemological changes of the Age of Enlightenment came a new culture and new concepts. The great Welsh tradition could only be maintained if it was shown that it could serve a great modern culture through an idiom thought to be pure – authentically its own rather than borrowed from neighbouring cultures, especially the increasingly dominant English. Welsh was positioned as having this potential because of its ability to adapt (or, to be adapted) to accommodate the new concepts. This stance is suggested in the title of a booklet published in 1771 by John Walters, *A Dissertation on the Welsh Language, Pointing out it's [sic] Antiquity, Copiousness, Grammatical Perfection, with Remarks on it's Poetry*. According to M. Morgan (2002: 146-7), 'copiousness' and 'grammatical perfection' referred to

“cyflawnder geirfa'r Gymraeg a gwyhder gramadeg [...] Clodforai ieithyddion y ddeunawfed ganrif y rhinwedd arbennig hon yn anad dim arall. Gellid ffurfio geiriau newydd o adnoddau brodorol y Gymraeg yn hawdd ac oherwydd hynny, gellid ei chadw'n iaith bur.”⁵⁴

Thus began the gap-filling tradition in Welsh lexicography, two of the most significant early elaboration lexicographers being John Walters and William Owen Pughe⁵⁵. Walters' and Pughe's dictionaries introduced a large number of new words, many of which are in use today⁵⁶. Some examples are given in Table

⁵⁴ 'the copiousness of the vocabulary of Welsh and the brilliance of its grammar [...] Eighteenth century linguists lauded this particular quality above all else. New words could easily be formed from the native resources of Welsh, and Welsh could therefore be kept a pure language.'

⁵⁵ They compiled the dictionaries *An English-Welsh Dictionary...* (Walters, 1794); and *A Dictionary of the Welsh Language...* (Pughe, 1793-1803). Morgan lists a number of other influential dictionary-makers who coined words, but considers Walters and Pughe to have been particularly significant.

⁵⁶ I shouldn't overemphasise Pughe's success as a coiner of words, however. He coined many words, but his neologisms were largely founded on his now strange ideas about the origins of language (see Jenkins (1997) for details). This and his attempt to revise an already fairly stable orthography made his coinings largely opaque and 'a total mystery to ordinary readers' (Jenkins, 1997: 392, my translation).

4.1.

Walters' coinings	English gloss	Pughe's coinings	English gloss
adloniant	entertainment	alaw	melody
adnodd	resource	amgueddfa	museum
ad-olygu	to revise	argymhelliad	recommendation
amaethyddiaeth	agriculture	arweinyddiaeth	leadership
arall-eirio	to paraphrase	awyren	aeroplane
arddegau	teens	calonogol	encouraging
arwerthiant	auction	cynrychioli	to represent
awyrgylch	atmosphere	cyntefig	primitive
braslun	sketch, outline	damcanu	to hypothesise
bytholrwydd	perpetuity	darlith	lecture
clytwaith	patchwork	delwedd	image
cofnodi	to record	diddorol	interesting
cydnabyddedig	recognised	dylunio	to design
cyfanwerthu	to sell wholesale	ffaith	fact
cyhoeddusrwydd	publicity	ffrwydro	to explode
cylchlythyr	circular (letter)	gorwel	horizon
cylchrediad gwaed	blood circulation	gwyddonol	scientific
gorbriso	to value too highly	nodwedd	feature
meddygfa	(doctor's) surgery	popeth	everything
poblogaeth	population	rhwystredigaeth	frustration

Table 4.1: A small selection of words first attested in the dictionaries of Walters and Pughe, taken from Morgan (2002: 163 and 183)

Although there are some differences between these early dictionaries and more recent lexical planning efforts, they seem to share the gap-filling motive as well as, in part, its ideological foundation. Morgan (2002: 181) writes:

“Un o brif amcanion John Walters oedd galluogi pobl i drafod pob math o bynciau yn y Gymraeg. I’r perwyl hwn, ceisiodd ofalu bod

geirfa'r Gymraeg yn datblygu ochr yn ochr â'r Saesneg a cheisiodd greu termau i gynrychioli'r cysyniadau newydd.”⁵⁷

That is, in the terms of the language planning literature, Walters was concerned with status planning. But this status planning was carried out in the service of an ideological construct, a pure form of Welsh that had to be constructed in opposition to English. It was not enough for Welsh speakers to use English borrowings to talk about all sorts of modern subjects in Welsh. The lexical resources had to be Welsh because anything less would suggest that Welsh was not copious (complete, whole, adequate) and lacked brilliance. I show below how this idea persists in current lexical planning endeavours. Note how the Welsh coinings in the list above are made up of mostly Welsh morphological resources: *bras-lun* (literally ‘rough picture’), *awyr-gylch* (‘air circle’ / ‘surrounding air/sky’), *awyr-en* (‘sky’ + diminutive nominal suffix), *pobl-og-aeth* (‘people’ + adjectival suffix + nominal suffix). This is not lexical planning through borrowing and standardised integration, but of word-formation using Welsh morphological resources⁵⁸. The ideological foundation of lexical planning in this period seems to be quite dominantly purist, an ideology which, I argue, has persisted (see below).

I have talked so far about the early stages of Welsh standardisation. But in how far can we talk of a standard Welsh and standard lexical resources in a historical sense? Considering that standardisation initially happened through the translation and dissemination of religious texts, it might be argued that what developed at this early stage was not so much a standard register but a religious one. However, the profound influence of the 1588 Bible on Welsh society,

⁵⁷ ‘One of John Walters’ principal objectives was to enable people to discuss all sorts of subjects in Welsh. To this end, he tried to ensure that Welsh vocabulary kept abreast of English vocabulary, and he sought to create terms to represent new concepts.’

⁵⁸ On the other hand, the study of language was, at the time, in its infancy. Especially due to Pughe’s now strange ideas about the origin of language (see note 56), his neologisms are often hard to characterise as being made up of ‘Welsh morphological resources’, as we would understand this today. Nonetheless, his neologisms were made up of what he interpreted to be monosyllabic meaningful constituents of Welsh origin. He thus positioned his neologisms as being symbolic of the inherent copiousness of Welsh. It is this symbolism that is of significance here, rather than objective etymological fact.

comparable to the influence of Luther's Bible on German society (J. Davies, 2007: 221), meant that it served as a model for Welsh in literature and poetry, journalism and education (D. G. Jones, 1988). Moreover, whilst it is unclear what exactly was the extent of knowledge of, competence in and acceptance of this religious/literary/standard register, it is notable that literacy arrived relatively early in Wales. By means of privately-funded schooling that had a religious and social advancement agenda, the majority of the population of Wales were literate, in Welsh, by the second half of the 18th Century, a rarity in Europe at the time (J. Davies, 2007: 280). It is largely held, then, that through religious texts in particular, but increasingly through other written genres, the Welsh population largely acquired knowledge of some form of standard register, at least in writing. What exactly was the influence of lexicography, it is not clear. But undoubtedly, Welsh speakers came into contact with the lexical resources used in the religious/literary/standard register, including the new words that were devised for the influx of new concepts that were a consequence of the rapidly changing culture, in particular during the 18th Century.

It is notable, however, that this register wasn't an administrative register, after the Act of Union of 1536 in effect ousted Welsh from state administration in favour of English, a situation that would last until the 1967 Welsh Language Act. As a result, two standard languages were used in Welsh-speaking Wales, endogenous (standard Welsh, used in literature, religion, etc.), and exogenous (standard English, used in public administration). The partial domain-restriction of the standard register meant that knowledge and acquisition of it was restricted to those that interacted with those domains, which entails sociolinguistic variation. Moreover, interaction with the domains of the standard languages used in Wales has not been static. With the contraction of its domains in the 20th Century, particularly the religious (D. D. Morgan, 2000), and the expansion of those domains where the exogenous standard was dominant, namely state administration and education⁵⁹, the Welsh standard register largely

⁵⁹ The 1870 Education Act ensured that English would become the official language of the British

lost its mode of acquisition and its authority. The rapid decline in the percentage of Welsh speakers in the population throughout the 20th Century likely further contributed to the decline in the number of people converse in the standard register.

It is also possible that the rigidity of its norms (or of those who imposed them) and its divergence from ordinary spoken styles was a confounding factor in the decline in number of people converse in the Welsh standard. Developed out of Bishop Morgan's 1588 translation of the Bible, which was itself based on the strict pan-Wales code developed in the bardic schools of the Middle Ages (G. E. Lewis, 1987), the standard register was esoteric, even after doing away with Pughe's strange influence. It was a craft to be learnt diligently and used skilfully, a "special kind of language to be learnt and mastered, and not tinkered with" (D. G. Jones, 1988)⁶⁰. B. Thomas & P. W. Thomas (1989: 6) cite Hugh Evans, who warned as early as 1931 that there was a danger "y byddai poeni am wneud camgymeriadau yn atal y Cymro cyffredin rhag mentro ysgrifennu yn *Gymraeg*"⁶¹.

The difficulty of the standard register and the loss of its authority in the 20th Century led to a debate concerning whether it should be modified. Some felt it was too far removed from spoken practices, drawing a comparison between 'book Welsh' and the 'living Welsh' of spoken practice (C. Davies, 1988; I. Williams, 1960). This debate culminated in the development of a teaching model called *Cymraeg Byw* ('living Welsh') (Welsh Joint Education Committee, 1967, 1970), a redoing of corpus planning (Clyne, 1997). Opponents of *Cymraeg Byw* argued for the vitality and authority of the standard register. For example, C. H.

state education system (C. H. Williams, 2000: 642), although some provisions were later made for the teaching of Welsh within state-funded schools (Evans, 2000).

⁶⁰ Here D. G. Jones is characterising how Morris-Jones (a harsh critic of Pughe's) would position standard/literary Welsh in his Eisteddfod adjudications.

⁶¹ 'that worrying about making mistakes would prevent the ordinary Welsh speaker from venturing to write in Welsh'.

Thomas (1967: 242) insisted that there already was a:

“iaith lafar safonol sy’n gymeradwy gan bob siaradwr Cymraeg, boed anllythrennog, boed hyddysg, iaith hysbys i bawb ac arferedig ganddynt hyd y mae gafael weithredol ganddynt arni.”⁶²

Such pronouncements rejecting *Cymraeg Byw* sought to reassert the authority of the standard register, an authority it seemed to be losing to that of the emergent Welsh-medium (pre-university) state education system, which eventually replaced the chapels as the principal agent through which Welsh literacy was taught. *Cymraeg Byw* does not relate directly to lexis, since it primarily focused on grammatical forms. However, the initiative is suggestive of changing norms and changing authority, something which also potentially impact on the acquisition and use of lexical resources.

Education is now one of the government’s cornerstone language planning policies (Welsh Government, 2012b)⁶³. Increasing numbers of pupils received Welsh-medium education in the 20th Century, including many who had not acquired Welsh at home. In 1965, there were 142 pupils in designated bilingual schools. By 1990, this figure had risen to 11,519 (Evans, 2000: 353)⁶⁴. M. C. Jones (1998) considers the effect of Welsh-medium education on language practices. In Rhosllannerchrugog, a community in north-east Wales, first-language school-age speakers who weren’t receiving Welsh-medium education evidenced a number of more traditional dialect features in their casual speech than did their counterparts who were receiving Welsh-medium education. Jones took this to be evidence that Welsh-medium education was having a standardising effect on casual, spoken styles.

⁶² ‘a standard spoken language acceptable to all Welsh speakers, whether they be illiterate, whether educated, a language which is familiar to all Welsh speakers and used by them if they are actively competent in it’.

⁶³ There is no obligation for parents to send their children to Welsh-medium or bilingual schools, although Welsh as a subject is compulsory (1st or 2nd language) up to 16 years.

⁶⁴ It is important to note that Welsh-medium education began as a grassroots movement, rather than being a government initiative.

There is also some concern over perceived changes in spoken practices towards more contact effects and simplification (E. L. Jones, 2002; A. R. Thomas, 1982), as well as the loss of Welsh lexical resources, termed *lexical erosion* (A. E. Jones, 1982, 1985). ‘Change and decay’ (A. R. Thomas, 1982) seems to be an ever-present concern, but one that is heightened by the realisation that second-language speakers don’t necessarily emerge from Welsh-medium education with first-language-like proficiency, or not with such proficiency that easily fits within a first-language perspective of what is ‘authentically’ Welsh. I found evidence in this respect in a survey investigating language attitudes towards Welsh speakers in south-east Wales (Robert, 2009). Second-language styles were typically more negatively evaluated than first-language styles, although there was a hierarchy of sorts of second-language styles, differentiated according to how correct respondents perceived them to be. More ‘correct’ second-language styles were rated more positively than less ‘correct’ styles, which typically evidenced more contact effects and simplification. There is a belief that these language practices are also being taken up by first-language Welsh speakers, possibly due to the low number of first-language pupils relative to second-language in Welsh-medium schools in some areas.

I have presented above some details on more recent sociolinguistic variation that is relevant to a discussion of standardisation. But to return to standardisation specifically, it is not entirely clear that we can talk about ‘standard Welsh’ at all, since it is unclear that there is a delineated set of language practices that are oriented to by a powerful group of Welsh speakers as befitting all occasions that is positioned by those in authority as *normal*. Certainly, there are ideas about what is ‘good’, ‘proper’, ‘rich’, ‘*graenus*’ (‘strong’ or ‘vigorous’, literally ‘grainy’), ‘*gloyw*’ (‘bright’/‘shining’) and what is ‘bad’, ‘wrong’, ‘*bratiog*’ (‘ragged’), but such ideas seem, in part, to be contextually sensitive. Coupland (2009) has argued this point in relation to English, highlighting that language attitudes research suggests that varieties have more complex profiles than ‘standard’ and ‘non-

standard', 'dialect' or a similar label, and that indexicality is culturally embedded. However, 'standard Welsh' as a social construct seems to be much less firmly consolidated than 'standard English', and this must be attributed to the coexistence of two standard languages in Wales, and the contraction of the domains of standard Welsh in the 20th Century. Nonetheless, the *idea* of standardness seems to be quite firmly rooted, perhaps as a result of extremely close cultural contact with English (Jaffe, 2003 makes a similar point in relation to Corsican and French).

4.3 LEXICAL VARIATION AND BORROWING

I have primarily focused above on standardisation in general terms, although I have given some detail on early lexical planning. What can we draw from the discussion specifically regarding lexis? Firstly we can say that lexical planning shows a preference for 'Welsh' forms, and that this has been a trend since the early days of dictionary-making. This, then, provides some scope for lexical variation, since the 'standard' language discourages borrowing, whereas in more vernacular practice Welsh speakers do borrow. It is not necessarily the case that planned words have come to displace borrowings, but that borrowings and planned words have come to exist side-by-side. This is particularly true since the use of English has persisted and increased in Wales. Lexical planning cannot displace borrowing, since all Welsh speakers are now bilingual. This results in what A. R. Thomas (1987) has called *doublets*, or what we might also call *lexical variants* (different lexical items used to denote roughly the same concept). We might say that two or more lexical variants make up a single variable, to borrow variationist sociolinguistic terminology. The variation in this case is contact-induced, since at least one of the variants for each variable is a borrowing. Note that there are cases of lexical variation in Welsh that do not involve language contact (e.g. regionally-differentiated words such as *llaeth/llefrith* ('milk')). Note also that it is not necessarily the case that all contact-induced variables include one borrowing and one *planned* word. Some of the 'Welsh' forms I list below date

back to at least the 13th Century⁶⁵, according to *GPC*, and are therefore quite likely not planned.

Some examples of contact-induced lexical variables by Thomas are *cerddoriaeth/miwsig* ('music') and *diffodd/rhoi allan* ('to extinguish'). In the case of the latter, the borrowing (*rhoi allan*) is a loan translation, the "combination of native morphemes in imitation of a foreign pattern" (Winford, 2003: 45). Other examples are: *ffodus/ffortunus* ('fortunate'), *defnyddio/iwso* ('to use'), *peryglus/danjerus* ('dangerous'), *hedfan/fflio* ('to fly'). We might also include cases where a borrowing is less integrated, such as *esgid/shoe*, *rhyngrwyd/internet*. Although some would contend that in such cases the borrowed variant is not 'Welsh' at all (in contrast to integrated/well-established variants), and hence there is no variation, I argue that this is a matter of perspective. There is no way of knowing exactly how individuals perceive these words, or whether drawing a line between them is meaningful at all. To assert that a word is not 'Welsh', and therefore does not count as a variant is arguably the assertion of a prescriptive norm rather than a statement of objective fact. It is part of the discourse of standardisation and purism.

I emphasise, however, that as resources for meaning-making, lexical items don't work alone as single bounded units of social meaning. Lexical items have the potential to be integrated, or not, through phonological, morphological, even syntactic and semantic means – through the word's 'constituent patterns', which can be treated differently by bilinguals (Haugen, 1972 [1950]). Winford's (2003) framework of lexical borrowing is useful for seeing the different types of borrowing and integrative processes that provide bilinguals with resources for meaning-making. I reproduce his framework in Table 4.2, with some of my own additions, and examples specifically from Welsh. I have taken these examples from my dataset, to ensure my examples are attestable.

⁶⁵ *GPC* lists dates of attestation. In the case of words whose date of first attestation is early, their use likely extends back much further.

I should note that I have only presented here the first part of Winford's model of lexical contact phenomena, namely that part dealing with what he calls *borrowings*. In his original model, Winford also has a further category, *native creations*, namely those linguistic resources that are created through formal or informal lexical planning, as a result of new concepts being introduced through language contact. I have not included this category here, since my main interest in this section is in lexical borrowings. However, it ought to be noted that Winford views 'native creations' – and thus the majority of the lexical planning endeavour for Welsh – as lexical contact phenomena.

Two terms need explaining here: importation and substitution, terms that Winford has borrowed from Haugen (1953, 1972 [1950]). *Importation* refers to the borrowing of a meaningful linguistic unit, namely a morpheme or a semantic meaning, or possibly a 'foreign pattern', in the case of calques. *Substitution* refers to the process of swapping sounds or morphemes for those in the borrowing language (Winford, 2003: 43), a process I refer to instead as *integration*. I argue also, however, that in some cases, syntactic integration is also possible. If we accept a definition of lexical item to include multi-word items (as I have discussed in Chapter 3), then the syntactic word-formation rules used to form multi-word items are also amenable to substitution/integration. In the case of Welsh and English, their noun-phrase word orders are in reverse: head-modifier in Welsh, and modifier-head in English. This provides Welsh/English bilinguals with a resource for integrating, or not, borrowed multiword items. Thus I have added the category *syntactic substitution* to "Pure" loanwords above (the quotation marks are Winford's), with the example *arthritis rheumatoid*.

Types	Processes involved	Examples from data
A Loanwords ⁶⁶ :		
1 “Pure” loanwords (Total morphemic importation of single or compound words)		
	Without substitution	<i>access, autoimmune</i>
	With phonemic substitution	<i>busnas</i> (‘business’), <i>consyrn</i> (‘concern’), <i>seis</i> (‘size’)
	Syntactic substitution ⁶⁷	<i>arthritis rheumatoid</i> (‘rheumatoid arthritis’)
	Semantic change	<i>job</i> (difficulty), <i>saff</i> (certain)
2 Loanblends (Combination of native and imported morphemes)		
2a Derivational blend	Imported stem + native affix	<i>supplyio</i> (to supply), <i>gwaetsiad</i> (to wait)
	Native stem + imported affix	No Welsh examples. Winford’s example: Jap. <i>ichigo-edo</i> (‘strawberry’ + <i>-ade</i>)
2b Compound blend	Imported stem + native stem	<i>sheet gwyrdd</i>
B Loanshifts (loan meanings):		
1 “Extensions” (Shifts in the semantics of a native word under influence from a foreign word)		
	a Phonological resemblance	No Welsh examples. Winford’s example: Am. Port. <i>humoroso</i> (‘humorous’)
	b Partial semantic resemblance	No Welsh examples. Winford’s example: Am. Port. <i>frio</i> (‘cold infection’)
2 Loan translations (calques)	Combination of native morphemes in imitation of foreign pattern	<i>stafell wely</i> (bedroom), <i>setio i fyny</i> (to set up), <i>sefyll allan</i> (to stand out)

Table 4.2: Winford’s framework of lexical borrowing (adapted from Winford, 2003: 45)

However, in terms of the potential for borrowing and integration process to provide bilinguals with resources for meaning-making, Winford’s framework has two problems. Firstly, interpreting whether or not a borrowing is integrated

⁶⁶ Winford’s use of *loanword* is not to be confused with Poplack’s (see Chapter 3), for whom *loanword* refers to well-established borrowing. For Winford, it would seem, *loanword* refers specifically to the fact that morphological content is near identical (if sometimes integrated phonologically and/or syntactically) to the lexical item as used in the borrowed language.

⁶⁷ This is my own addition to the framework

must rest on assumptions about the shape of the word in the borrowed language. This is not always clear, in particular in the case of pronunciation. In some cases, any possible pronunciation differences would be so small as to be almost inaudible (e.g. the integrated pronunciation /'dɔkdɔr/ in spontaneous speech is not easily distinguished from /'dɔktə/). In others, classifying a borrowing as integrated depends on our interpretation of how the word would/should be pronounced if it weren't integrated, which is a matter of normativity. Since accent is variable, there is no one pronunciation to compare with. For example, *stroc* ('stroke'), pronounced /strɔ:k/, can be considered integrated, and yet it would normally be realised the same way in a number of British English accents, including a number in Wales. *Stroc*, then, is integrated compared with /strəʊk/, but not compared with a number of Welsh English accents. This is not to say that there is not potential social meaning difference between /strɔ:k/ and /strəʊk/, only social meaning does not necessarily map neatly onto 'Welsh' and 'English'.

Secondly, Winford claims that *importation* and *substitution* are "a dichotomy based on the presence or absence of foreignness markers" (43). This is only partly true. There are some words that, whilst they may not be integrated on the word level in any way, still contain no foreignness markers. In terms of the potential of lexical items and their 'constituent patterns' to signify socially, I believe this is significant. For example, in the table above, I have included as examples of "pure" loanwords without substitution both *access* and *autoimmune*. Within Winford's framework, then, both these words are treated equally. However, they differ markedly, within the Welsh/English bilingual context, in their openness to integration and in the degree they might be interpreted as foreign. There is nothing that can be done to *access* to make it sound more Welsh, whereas *autoimmune* is a compound that includes two morphemes, at least one of which could be substituted for a more 'Welsh' equivalent (*auto* for *hunan*, as we also find in the *Syrjeri* data). It is, then, much more open to being interpreted as foreign than *access* is.

I argue, then, that whilst Winford's framework is useful for understanding borrowing processes and labelling borrowings, we also need to consider how markedly 'foreign' borrowings are (or might be perceived to be), according to their constituent morphemes, phonemes, etc., rather than merely focusing on how markedly integrated they are. After all, if purism is an ideology that responds and seeks to restrict the effects of contact phenomena, then it is likely that words that are more markedly foreign will be more prevalent than others in terms of social meaning. However, this does not mean to say that unmarked, non-integratable borrowings, such as *access*, are not interpretable as borrowings. I contend that it is not only a word's constituent parts that might cause a speaker to infer that it is a borrowing. We must also consider the speaker's history, their knowledge of the borrowed language, degree of linguistic expertise, as well as wider normative assumptions.

The availability of contrasting lexical forms and their 'constituent patterns', then, opens up the scope for variable social meaning. A. R. Thomas (1987) asserts that in his doublets (above), 'indigenous' forms are typically seen as 'standard'/formal and borrowed forms as 'non-standard'/colloquial. It is possible to see how this may result from differentiated practice and engagement, or not, with those institutions and activities where planned Welsh has traditionally been acquired and used. Jilg's (2003) study of the lexical competence of Welsh speakers in Blaenau Ffestiniog, a town in north-west Wales, is insightful in this regard. Jilg set out to assess Welsh speakers' knowledge of Welsh words vis-à-vis English, for long-standing and more contemporary concepts. He compared competence across the two most salient social networks that he'd identified in the town: *Pobly Pethau* ('the 'pethau' people')⁶⁸ and *Pobly Dafarn* ('the people of the pub'). Although Jilg doesn't explain why exactly he chose these specific labels for his social networks, arguably, the former refers to those who more typically engage with the domains of the Welsh standard register and the latter refers to

⁶⁸ 'Y pethau' is a culturally-specific term, literally meaning 'the things' but referring fairly obliquely to Welsh-language cultural institutions and activities, such as the Eisteddfod, drama and poetry.

those who do not. Using visual stimuli, Jilg elicited lexical items for a number of concepts (e.g. relating to machines, household items, job titles, etc.). He found that *Pobly Pethau* were more likely to respond with Welsh forms (rather than English) than were *Pobly Dafarn*, in particular in the case of ‘rare’ Welsh words. The vocabulary of *Pobly Pethau*, then, is more ‘Welsh’ than that of *Pobly Dafarn*. One of the extralinguistic differences he found between his groupings was that *Pobly Pethau* tended to have attained a higher level of education than *Pobly Dafarn*. Jilg also notes that younger speakers responded more frequently with Welsh forms than did older, which he explains as an effect of increasing Welsh-medium education (where older generations are more likely to have received formal education through the medium of English). Jilg’s study potentially points to meanings that might be indexed by the use of borrowings or ‘Welsh’ words. As one of his participants noted:

“Ma’ rhaid iti, yng Nghymru ... addasu dy iaith ar gyfer bobl y pethe a bobl ddim y pethe neu ma’ ’na beryg ofnadwy iti gael dy ystyried, ar gam, o ryw snobiaeth ieithyddol.”⁶⁹ (2003: 131)

There is a potential association between language and class, or status group, which, from a certain perspective, can lead to the negative evaluation of such linguistic forms. Nonetheless, the study doesn’t shed light on how lexical resources are deployed, or on the indexical value or social meaning achieved by borrowed/Welsh lexis in practice. As I have argued previously (see Chapter 3), linguistic forms do not have fixed social meanings.

From the standard ideological perspective, however, it is clear that borrowings are typically negatively evaluated. Indeed, D. G. Jones (1988: 156) singles out “lightly touched-up borrowings” as the only category of words that are “unacceptable in literary [standard] Welsh”. C. Williams’ (1999: 3) language

⁶⁹ ‘In Wales you have to ... adapt your language for the pethe people and not the pethe people or you risk being considered, wrongfully, of some linguistic snobbery.’

guidelines state that writing Welsh clearly does not mean using “mwy o eiriau Saesneg, bratiaith a.y.y.b” (‘more English words, ragged language, etc.’) but rather, in bolded type, “Symleiddio gan gadw *urddas* y Gymraeg” (‘simplification whilst retaining the *dignity* of Welsh’, my emphasis). Using English borrowings, then, is seen to be on the same level as *bratiaith* and takes away from the dignity of Welsh – a statement that reflects a very similar ideology to that found in my discussion of Walters and Pughe in the previous section.

Through a wider frame, we can view lexical planning and the dislike of borrowings to pertain to the position of Welsh as minoritised language vis-à-vis English. This is also something that seems to have persisted since the early days of lexical planning. Of Pughe, Morgan writes that his motive was partly ‘the love of his country’ (M. Morgan, 2002: 168), to “preserve the remains of a language of an ancient nation” (Pughe, cited in M. Morgan, 2002: 168), but also a need to defend Welsh against the unfavourable attitudes of others, such as his friends in London, who scorned him for “What they called so unprofitable a sacrifice of time as the collecting together the words of a nearly expiring language” (Pughe, cited in M. Morgan, 2002: 168). Thus the gap-filling tradition could be seen as a defence of the historical canon of Welsh, which, in the London context where Pughe grew up, encountered the hegemonic discourse of English imperialism. Here there are striking similarities with more contemporary lexical planning, as we saw in the quote by J. E. C. Williams in Chapter 3 (p.66). Williams claims that the lack of ‘adequate’ linguistic resources in Welsh has in the past, through a process of widespread borrowing, the use of ‘patois’ and a feeling of linguistic inferiority, led to Welsh speakers abandoning ‘their native tongue’, i.e. language shift. Williams doesn’t talk about English imperialism per se, but he is referring to the minority status of Welsh vis-à-vis English and to the hegemonic discourse that is seen to have caused the minoritisation of Welsh. Conversely, lexical planning is positioned as part of the contemporary revitalisation and counter-hegemonic project.

C. H. Thomas (1982) makes a similar case, in particular that the disadvantaged position of Welsh and the lack of official recognition for Welsh (at the time) had hampered the development of a full register range in Welsh. That meant that “scientists or doctors, like public servants, were usually incapable of handling their own special fields in their own language. They lapsed either entirely into English or into a mixed language in which all the special terminology was in English” (C. H. Thomas, 1982: 95), a situation that was considered unacceptable by the ‘Welsh-speaking community’, according to Thomas who, as an academic, takes on the problematic role of speaking on behalf of ‘the Welsh speaking community’. Thomas considered lexical planning, especially in specific subject areas, to be the answer. But the ‘unacceptability’ of this situation is a matter of ideology – it pertains to the ‘inappropriateness’ of the use of borrowings – a point made clear when Thomas writes further about ‘mixed language’ (borrowing in specialist and other talk):

“There is only a narrow margin between this kind of mixed language and the sort that results from a breakdown of a speaker’s control of Welsh brought about by the deficiencies of the system of education. Because both diglossia and mixed language are associated in Welsh minds with the disadvantaged position of Welsh, both are considered to be unacceptable phenomena that must not be permitted to be lasting.” (1982: 101)

Borrowing is seen to be symptomatic of a social system that disadvantages the viability of Welsh-language culture, and of the individual speaker’s personal lack of control over their linguistic capacities. I refer the reader here to what I noted above about standard language ideology in the past, where the standard register was seen to be a craft to be learnt diligently and used skilfully. Use of borrowings, in Thomas’ view, signifies a lack of linguistic competence – caused, in part, by a deficient education system. Moreover, this is something that needs to be brought under control.

Thomas' comments, however, were made over 30 years ago. Are the same ideas and assumptions still in circulation today? I wrote in the previous section that changes in demography and the social practices of Welsh speakers led to changes in the sociolinguistic profiles of Welsh speakers – in the linguistic resources in their repertoires. Moreover, social changes challenged the norms and authority that had previously upheld the authority of the standard register. It seems to be the case that there are increasing voices urging for more leniency and tolerance in matters of language (arguably, my own work is part of such a trend). For example, during a symposium on language planning that I attended in north-west Wales⁷⁰, a number of participants in breakaway discussion groups noted that Welsh speakers ought to be more tolerant of the linguistic practices of others, in order to encourage those less confident in their use of Welsh to use the language in situations where they might otherwise use English (like in the hypothetical situation outlined by Thomas above).

In a more official capacity, in his report on Welsh-language provision in the *National Health Service (NHS)* in Wales, Misell (2000) discusses the need felt by a number of healthcare practitioners for lexical planning in healthcare, to forge a “common language to enable consistent meanings to be attributed to words, terms and data” (Welsh Office 1998, cited in Misell, 2000: 51). This common language is said to be needed for the dissemination of healthcare information to patients (or ‘healthcare consumers’). However, Misell finds this monologic conceptualisation of language – an authoritative code where all meanings are fixed prior to interaction, and where there is little room for responsivity (Lotman 1988, cited in Wells, 2007) – problematic in the case of Welsh, because of mixed competencies, confidence levels and access to planned language. Misell notes that, in interviews conducted with healthcare consumers, a number of interviewees “insisted quite definitely that they found it easier to discuss health-

⁷⁰ *Tu Hwnt i'r Dosbarth: Dyfodol Cynllunio Ieithyddol* ('beyond the classroom: the future of language planning), Bangor, March 2013

related matters in English than in Welsh because they didn't have sufficient vocabulary to talk about such a subject in Welsh" (2000: 52). What is notable here is not so much what Misell's interviewees said, but Misell's own response. Taking a consumerist standpoint (acting in the interests of healthcare users/consumers), Misell argues that the priority should not be "the maintenance of the Welsh language in all its spotless purity; rather to open the way for consumers to communicate in their preferred language and at the level at which they are most comfortable" (2000: 53). It is possible, then, that there is a shift towards more tolerance and accommodation to variable language practices. Nonetheless, it is impossible to say from the examples I have given whether there is a widespread shift, or merely a small number of voices.

4.4 PLANNERS, PRODUCTS AND PROCEDURES

I have given some detail on historical lexical planning in section 4.2. In this section I consider more recent lexical planning that might have a more direct impact on current language practices, covering some of the linguistic tools developed for use by Welsh speakers, and the administrative framework that might have an impact on the take-up of contemporary lexical planning.

In the previous chapter, I discussed some of the codification texts I would consult in my analysis. The texts I mention (*Geiriadur yr Academi*, *Termiadur* and *Cysgeir*) are amongst the most significant products of recent lexical planning. There are a small number of other general lexical codification texts, most significantly *Geiriadur Prifysgol Cymru* ('the University of Wales dictionary', henceforth *GPC*) (University of Wales Press, 1950-2002), the only substantial Welsh dictionary, which was published in instalments between 1950 and 2002 (and is now being revised). Being a historical dictionary, it is much less prescriptive than most other Welsh lexical codification texts, although still contributes to the codex of Welsh – a definition of a bounded entity called 'Welsh' which perpetuates the idea that this bounded entity is an objective reality. Although it is the closest Welsh equivalent to the authoritative *Oxford English Dictionary*, the sheer size

(and consequently, price) of the dictionary has meant that its circulation is more restricted than other lexical codification texts, although dictionary staff have been working in recent years on an online, concise, searchable version⁷¹. In terms of lexical elaboration, its potential is restricted by its previous bit-by-bit publication – recent neologisms beginning with a letter nearer the beginning of the alphabet (such as *e-bost* ('e-mail')) are not included, since they had not been coined at the time of publication. Staff have started working on updates to the earlier instalments, but understandably progress is slow. Simply, the work on the dictionary is too slow, and financial resources too scarce, to keep abreast of language change.

Apart from large, general lexical codification texts, a number of subject-specific texts have been published since the middle of the 20th Century. The focus on subject-specific lexis went hand-in-hand with the establishment of Welsh-medium education during the same time (Prys, 2006), and most of these texts were published by two education bodies, the *Welsh Joint Education Committee* (WJEC) and the *University of Wales Press* (UWP) (the latter works were carried out by the *Board of Celtic Studies* established by the University of Wales). These efforts involved compiling bilingual English-Welsh glossaries in particular subject areas, for example music (UWP, 1978; WJEC, 1992a), history (UWP, 1959; WJEC, 1987), mathematics (UWP, 1957) and physics (UWP, 1965; WJEC, 1983). Such glossaries were designed for use in the compilation and use of teaching materials and examinations.

The 1993 Welsh Language Act, the act that placed an obligation on public bodies operating in Wales to provide some level of Welsh-language service, led to further perceived demand for 'Welsh' terms, and a number of other lexical codification texts have been published, in varied subject areas. These include, for example, finance (Prys, 2000a), healthcare and associated fields (Prys, 2002b;

⁷¹ According to their website, http://www.cymru.ac.uk/geiriadur/gpc_pdfs.htm#DANGOSEIRIAU last accessed 20/08/2013.

Prys & Williams, 2001), law (Léwis, 2003), local government (D. G. Lewis, 1996), the creative industries (Ffrancon, Dafydd, Prys, & Davies, 2008) and sports (Iorwerth, James, & Williams, 2007). Many of the codification texts published in the last ten years or more are made available in digitised form (sometimes exclusively). As I have previously suggested, the aims of these glossaries typically differ: some are meant to be used in education contexts, some for facilitating the use of Welsh by specialist practitioners, some seemingly targeting leisure activities, some public administration, etc. These examples show the diversity of subject areas that are planned for. Some are not subject-specific, but they do centre on broader fields (for example, *TermCymru*⁷², a database of the terms used by the Assembly's translation service, made available online as a searchable database in 2003, contains vocabulary related to the work of the Assembly, diverse in scope, yet generally related to public administration). I have listed here only some examples of the various lexical codification texts published since the 1950s. There is a bibliography of codification texts up to 1995 in Welsh Language Board (1995) and also in Prys, Jones & Emlyn (1995). There was no updated list at the time of writing, but there is a selected list of work undertaken by the Terminology Centre at Bangor University on their website⁷³.

We see from these examples that lexical planning has typically been carried out by a number of practitioners. There is no Welsh language academy⁷⁴, a central, state-sponsored body dedicated to codifying a national language, “endowed with an aura of astonishing power to accept or deny the legitimacy of any given word” (Stavans, 2003: 31) such as the *Real Academia Española de la Lengua*. There have, however, been some tentative moves in the direction of the official centralisation of lexical codification and authorisation, but nothing concrete has emerged. For example, B. P. Jones (1988: 179) notes that “the Hughes Parry Report [on the

⁷² www.termcymru.wales.gov.uk

⁷³ www.bangor.ac.uk/canolfanbedwyr/termau.php.en?%20onclick

⁷⁴ The *Academi Gymreig* / *The Welsh Academy*, who published *Geiriadur yr Academi*, is not primarily a language academy, but describes itself as the national Society of Writers in Wales www.literaturewales.org/the-welsh-academy Accessed on 02/04/2013. Its only corpus planning work to date has been the dictionary, and it does not seem that it has plans to produce any more in the near future (D. G. Jones, 2013).

Legal Status of the Welsh Language, 1965] emphasized the importance it attached to entrusting the responsibility for selecting and coining suitable Welsh terms to a *single body*” (my emphasis), but no provisions were made for such a body in the 1967 Welsh Language Act that followed the report. In 1995, the Welsh Language Board published a report by the *Panel for Official Welsh* (Welsh Language Board, 1995) which recommended setting up a *Department for Welsh Language Standards*, which would be responsible for standardising terms for public administration. The department never came to light, but the Board sought to initiate lexical planning work under its own remit, commissioning a pilot project, *Geirfa Deddfwriaeth Priffyrdd / Highways Legislation Terms* (Prys, 1998) and accompanying guidelines in the form of a report on the project (Prys & Jones, 1998a), and later commissioning updated and more comprehensive guidelines (Prys & Jones, 2007). In 2008 it was working towards setting up a *Canolfan Safoni Termau Genedlaethol/National Centre for the Standardisation of Welsh Terminology*⁷⁵ (Welsh Language Board, 2008). However, when the Welsh Language Board was abolished in 2012, the Centre had not been established (the Board having not secured additional funding (Welsh Language Board, 2009b)), and it is unclear whether the Welsh Language Commissioner, who has taken over responsibility for the Board’s work in lexical planning, will continue to press ahead with plans.

It is probably fair to say that, although the Welsh Government has driven forward work in lexical planning to a certain extent, so far through the Welsh Language Board, lexical planning isn’t high on its list of priorities. We see, for example, that *Iaith Pawb*⁷⁶ dedicated only half a page (out of a total of 58) to ‘language tools’, and a mere two bullet points to ‘standardising terminology’ (Welsh Assembly Government, 2003: 49). In its more recent language planning strategy, *Iaith Fyw: Iaith Byw*, the Government notes only that it will “ensure that the language can

⁷⁵ This is not the *Centre for the Standardization of Welsh Terminology* at Bangor University.

⁷⁶ *Iaith Pawb* (‘everybody’s language’) was the Welsh Government’s ‘National Action Plan for a Bilingual Wales’. It was superseded by *Iaith Fyw: Iaith Byw* (‘a living language: the language of living’) (Welsh Government, 2012b).

benefit from a standardised online Welsh language dictionary, based on the significant investment it has already committed to this project over time” (Welsh Government, 2012b: 49). The accompanying action plan for 2013-14 states no more than that the Government will “ask the Welsh Language Commissioner to coordinate the standardisation of Welsh language terminology and place names” (Welsh Government, 2013: 19), a vague statement that makes no commitment to driving forward a lexical planning agenda. We are unlikely, then, to see any serious efforts at official centralisation in the near future.

Centralisation of codification work and authority has been sought by other means, primarily through efforts by practitioners to collate their lexical codification texts in a single database, accessible and searchable through a digital interface; specifically by the Welsh Language Board and the Terminology Centre at Bangor University. The Welsh Language Board launched a *Welsh National Database of Terms* in 2006 (hosted by the Terminology Centre at www.e-gymraeg.co.uk/bwrdd-yr-iaith/termau/Default.aspx), a requirement placed on the Board by *Iaith Pawb* (Welsh Assembly Government, 2003: 49). Recently, the Terminology Centre has collated the majority of its glossaries and those of “approved partners”⁷⁷ in a searchable, online database, *Porth Termau*⁷⁸ (‘terms portal’), and in the *Ap Geiriaduron* (‘dictionaries app’) for Android, iPad and iPhone, but had undertaken centralising work previously, most notably *Cysgeir* (see Chapter 3). The Terminology Centre has taken advantage of technological developments and employs in-house IT expertise to enhance its centralisation and dissemination work.

Whilst integrated online searchable databases undoubtedly have a practical benefit (it is much easier for the user to refer to a single source, rather than to consult numerous codification texts), the effect seems also to be to vest those sources more salient and more regularly consulted with more authority. Where

⁷⁷ www.termau.org/index.php/home/?lang=en Last accessed 20/08/2013.

⁷⁸ www.termau.org Accessed on 02/04/2013.

the objective is the normalisation of the planned language, it isn't surprising that authority is an issue if different planners make different recommendations, as we see when Prys, head of the Language Technologies Unit at Bangor University in which the Terminology Centre is based, bemoans the working methods and some of the decisions of the Welsh Language Board's terminology standardisation unit that are taken "without regard to standardization decisions taken by other authorities" (Prys, 2007b: 117), such as herself. Prys presents the Welsh Language Board as wrong here, and what this achieves is to take away legitimacy from their work, which conflicts with her own, which, in turn, is presented as legitimate.

Other ways the Terminology Centre seeks to ensure that its planned language is used is to follow an approach that they see as bottom-up, demand-led, which

"enables a wide spectrum of stake-holders to take part in terminology projects. Public servants in charge of their organisation's language policy, subject specialists, professional terminologists, translators, bilingual administrators, and educators have all had a part in formulating Welsh terms, and therefore feel ownership for those terms, and an interest in using them." (Prys, 2007b: 116)

In theory, working with what Prys calls 'stake-holders' ought to ensure better dissemination and take-up of the planned language. Following the example set by the Terminology Centre, the Welsh Language Board also started working more directly with specific organisations and partners directly involved in the fields planned for. The aim was: to better understand the need for planning; to secure the cooperation of subject specialists; to ensure users' feeling of ownership for planned language; and to aid the dissemination of the planned language (Welsh Language Board, personal communication).

However, dissemination and take-up of planned language are arguably most likely to be secured where there is an element of compulsion. For example, in Quebec, Law 101 (1977) made the use of lexis codified in the *Official Gazette of Quebec* mandatory in written materials published by government, in contracts to which the government was party, in education materials, and in research works approved by the Minister of Education (Weinstein, 1989). There is no such legal provision in Wales. However, organisations may insist on the use of planned words from a particular source in their own publications. In 1998, the *Qualifications Curriculum and Assessment Authority for Wales* (known as ACCAC, its Welsh acronym) published *Y Termiadur Ysgol* (Prys & Jones, 1998b), a dictionary of terms for schools compiled by the Terminology Centre⁷⁹. This is the clearest example of planned Welsh lexis that is almost certain to be disseminated. The dictionary is described on its website as⁸⁰:

“Sponsored by the Welsh Government, it provides standardized terminology for the field of education. These are the terms to be used in Welsh medium exams and assessments and in resources of all kinds for teachers and students.”

Moreover, considering the power and authority of the institution of state education, its impact is likely to reach beyond the specific context of schooling.

Finally in this section, I want to consider how planners go about codifying lexis and coining neologisms. Prys & Jones (2007) is one of the only lexical policy documents available outlining guideline procedures for lexical planning for Welsh. The guidelines were commissioned by the Welsh Language Board and the Welsh Government, and they cover a multitude of issues, including project management and ways of gathering terms to be standardised. They also detail the methods and principles, based on standards of the *International Organization*

⁷⁹ The Centre later published *Y Termiadur*. Both have now been merged into *Y Termiadur Addysg* ('dictionary of terms for education').

⁸⁰ www.termiaduraddysg.org/?lang=en Last accessed 20/08/2013.

for Standardization (ISO), that the authors follow in the formation of terms, which I focus on here. Note that although the guidelines are informed by the Terminology literature, the methods of term formation are similar to those B. P. Jones (1988: 178) claims had been ‘pioneered’ earlier by the Board of Celtic Studies. I would expect contemporary lexicographers to follow similar methods.

Three methods are typically used for forming new terms: the use of pre-existing forms for new concepts (semantic extension), creating new forms through derivation and compounding (word-formation), and borrowing. Prys & Jones (2007) offer an example of semantic extension (here through metaphor) in English, namely the extension of *virus* in its medical sense to refer to a computer *virus*. Often Welsh words follow the same pattern of semantic extension as English translation equivalents, such as the extension of *cawod* (‘(rain) shower’) to designate a shower used for washing in a modern bathroom⁸¹. An example of word-formation through derivation given by Prys & Jones is the verb *blaenoriaethu* (‘prioritise’), formed by adding the verbal suffix *-u* to the noun *blaenoriaeth* (‘priority’). An example of compounding given is *gweithgynhyrchu* (‘to manufacture’), formed from *gwaith* (‘work’) and *cynhyrchu* (‘to produce’). They also include acronyms as examples of word-formation, although they note that acronyms are typically not ‘popular’ in Welsh. Again, words formed through word-formation can follow patterns similar to English, such as *cynllun busnes* (‘business plan’) given in *Cysgeir*. Words such as these, however, can also be classified as borrowings, since they are loan translations. Other than loan translations, borrowings used are described as ‘direct borrowings’, such as *system*, found in *Cysgeir*. However, Prys & Jones (2007: 35) note that “there is a presumption in favour of native forms over foreign borrowings”. B. P. Jones (1988) notes that borrowings are typically not used when they pose orthographic problems. There is a sense in which all words ‘created’ are

⁸¹ The earliest attestation of *shower* in the latter sense in the Oxford English Dictionary is from 1873 (Oxford University Press, 2013), whereas the history of the first sense extends back to Old English. *GPC* only lists the earliest sense of *cawod* (first attestation from the 12th Century). The ‘c’ section of *GPC* was published between 1950 and 1967, which suggests that, at the time, *cawod* was not widely used in its bathroom sense.

borrowings. In his framework of lexical contact phenomena, Winford (2003) includes the category *native creations*, that is, the use of native resources (formal and semantic), or a combination of native and borrowed resources, to express foreign concepts. Notwithstanding the difficulty of knowing whether a concept is foreign or not in a context where there is near absolute cultural integration between the worlds of both languages in question, lexical planning in Wales is typically not in the business of coining neologisms, but of translating them.

Interestingly, Prys & Jones' (2007) guidelines don't give any indication as to whether and how borrowings should be integrated, or of the differentiated social meaning that can attach to differently integrated borrowings. They detail what they call 'principles of term formation', which appeal to: linguistic 'correctness' (naming specific codification texts as authorities); term transparency (semantic meaning should be derivable from the designation); consistency of the terminological system (monosemy and univocity (see Chapter 2)); semantic 'appropriateness' (avoiding misleading designations, "unfortunate tendencies or negative connotations" (2007: 39)); conciseness; and derivability (terms should be able to spawn derivations)⁸². I suspect they view the integration of borrowings as a matter of 'linguistic correctness' and delegate decisions (in theory) to the codification texts mentioned, or to intuition. If, as I argue, the bilingual's linguistic resources are all open to signifying social meaning through indexical processes, the integration of borrowings is of central importance to Welsh lexical planners.

4.5 THE MEDIA CONTEXT

S4C was established by the UK Broadcasting Acts of 1981 and 1982, and at the time the data was broadcast, S4C was primarily funded through tax revenue. S4C

⁸² Whilst Prys & Jones stress that, in theory, none of the principles listed is given priority over the others, with candidate terms scored on the total number of principles they satisfy, in practice derivability has proven important.

is responsible for providing a Welsh-language television service⁸³, and it is the only television channel of any kind (public or commercial) that does so. The channel owes its existence to the mobilisation of Welsh speakers in response to rapid language shift and cultural change in the latter half of the 20th Century. The establishment of S4C was resolutely fought for by language activists who sought to counter what they saw, rightly or wrongly, as the hegemonic discourse that was propagated through the “machine for brainwashing [the Welshman’s] children out of their language and culture”, causing “the erosion and the decline of our national language” (*Urdd Gobaith Cymru* statement to the Annan Panel, 1974, cited in Smith, 2000: 308)⁸⁴. S4C still serves the ends of this movement to some extent today⁸⁵, even if it faces pressures and challenges from other directions.

As well as being a fundamental part of a status and acquisition planning agenda, S4C was anticipated by many as an organ of corpus policy and planning, particularly as a means of consolidating the legitimacy of standard Welsh, which had largely lost its domains of reproduction. Smith (2000: 327) writes, apparently in sympathy with this view:

“Câi darlledwyr yn aml eu hystyried yn hyrwyddwyr Cymraeg
safonol ac nid oes amheuaeth na fu bwletinau newyddion a
chyfraniadau a sgriptiwyd yn dda yn fodd i nifer mawr o wrandawyr
gynefino â dull mynegiant a oedd yn parchu safonau traddodiadol a
dod yn gyfarwydd â geirfa newydd a weddai i anghenion bywyd

⁸³ Initially this meant a bilingual service, with Welsh-language content televised during peak hours, and English content (from Channel 4) at other times. Since 2010, with the switch to digital television, all of S4C’s television content has been nominally ‘Welsh-language’.

⁸⁴ Literally meaning the ‘guild of hope for Wales’, *Urdd Gobaith Cymru* organises Welsh-language youth events across Wales. It is more commonly referred to as *yr Urdd* (‘the Guild’).

⁸⁵ E. Williams (2009: 71) notes that S4C “has contributed to constructing the “historical we” in Wales”. He refers, by example, to television series *Y Tywysogion* and *Taith yr Iaith* (on the history of Welsh princes and the Welsh language, respectively), both shown in 2007.

modern.”⁸⁶

The Welsh Office, the UK Government’s arm in Wales, even commissioned a report circa 1985 that considered how best to study the effect of S4C’s language output on ‘the extended pattern of use’ and ‘the linguistic characteristics of the language over time’ (G. Williams & Thomas, n.d.). Williams & Thomas rationalise a focus on language use and style thus:

“S4C [...] can have a distinct effect upon the extent of the use of the language among existing speakers as well as upon the style of the use. This can follow from the status that accrues to a language through its domain extension, and/or through the increased tendency for speakers to familiarise themselves with use of Welsh in certain contexts through the media, that is, [sic] language domains can be standardised through their institutionalisation and legitimisation, with the media playing a central role in these processes. Part of these processes involve [sic] the direct effect upon language, for example through vocabulary use (lexification) and selective forms of standard vocabulary (hypercorrection).” (G. Williams & Thomas, n.d.: 25-26)

Nonetheless, S4C has not since taken a prominent role as an organ of dissemination and implementation of lexical planning. For example, in listing those that would be targeted in the marketing of the glossaries of the *North Wales Health Care Standardization Terminology Group* (see Chapter 5), Prys (2007a) does not list S4C or any media organisation.

This may be because some inherent tensions in S4C’s institutional priorities,

⁸⁶ ‘Broadcasters were often considered to be promoters of standard Welsh and there is no doubt that news bulletins and contributions that were scripted well were a means for a large number of listeners to familiarise themselves with a form of expression that respected the traditional standards and with new vocabulary that were fit for the needs of modern life.’

namely a tension between the purist element of the standard ideology and the reality of language variation that S4C (and *Radio Cymru*) had to (or wanted to) attend to. Smith (2000: 327) notes, for example, that the use of on-the-spot recording facilities and unscripted interviews, where S4C wished to embed the voice of the man-in-the-street or expert, led to what he calls “iaith lafar hynod fratiog” (‘incredibly ragged speech’)⁸⁷. This was particularly true in interviews with subject specialists whose command of Welsh was ‘deficient’. The reality was, if S4C wanted to embed the voice of the man-in-the-street or the expert within the voice of S4C, it couldn’t depend on a ready supply of idealised Welsh speakers. There were not enough people with the desired language competencies. S4C faced a related problem in making its programming accessible. Research commissioned by S4C in 1992, and cited in M. C. Jones (1998)⁸⁸, showed that S4C’s news broadcasts were not “conveyed successfully to the average viewer” on account of the ‘academic language’ and ‘standardized terminology’ that was used (M. C. Jones, 1998: 274). Prior complaints had been made that S4C’s news programming was difficult to understand due to the ‘highbrow’ language used. As a result, S4C made changes to its news programming, seeking to accommodate the ‘average’ and the ‘high brow’ audiences through different broadcasts of different formats and using different styles of language.

However, where S4C did seek to accommodate viewers, it was criticised. For example, the soap opera *Dinas* (‘city’), in seeking to portray the lives of Welsh speakers in Cardiff in the 1980s, represented the use of some bilingual practices. S4C was criticised in the Welsh-language press, some examples of which are quoted by Ball et al. (1988: 191, their translation):

⁸⁷ *Bratiog* is an adjective which shares its root morpheme with the compound noun *bratiaith* (loosely translated as ‘ragged language’). *Bratiaith* is a highly negative term often used to represent speech, which typically incorporates many influences from English, as sub-standard.

⁸⁸ Jones notes that it was not possible to cite the reference of the research, due to S4C’s confidentiality rules.

‘Why, for goodness sake, is it necessary to answer the phone in English, or to speak ‘the slender language’⁸⁹ every time when discussing business? [...] I am amazed that authors, producers and actors – all claiming that they are good Welshmen – are willing to insult our language like this.’

‘The likelihood is that we shall have to suffer some programmes – the fewer the better I say! – which will contain sub-standard Welsh and a jumbled language to please those people who have not learnt at school what constitutes correct language.’

These tensions in S4C’s institutional priorities are still apparent today. In October 2010, S4C held a ‘Viewers’ Evening’, where a panel consisting of S4C Authority members and high-position staff answered viewers’ questions about various aspects of S4C’s service⁹⁰. One question was, in effect, a complaint about the language used on S4C, in particular the use of ‘English idioms translated literally’. The other called on S4C to start a campaign to raise the standard of spoken Welsh across Wales, specifically targeting the use of English words (e.g. *so*, *actually*, *(be)cause*). However, vox pops presented during the course of the programme showed mixed attitudes towards the current use of language on S4C. Some shared the same views expressed in the questions, some were happy with the language used, some spoke positively of the variation and contextual sensitivity represented on S4C, and some complained that they didn’t understand the ‘big words’ or regional varieties (other than their own). No doubt S4C carefully selected which vox pops it would show, purposefully representing all views in order to deflect criticism: it cannot please everybody. But perhaps it is a fair defence. It is probable that S4C will always face these tensions as long as there is only one channel that televises Welsh-medium content.

⁸⁹ Probably Ball et al. are here translating ‘yr iaith fain’, an expression commonly used to refer to English which I have always taken to be pejorative, although I have heard others disagree.

⁹⁰ Viewers had previously been invited to send in questions, and these had been sifted through by S4C, who selected which questions would be addressed.

S4C has an overtly formulated policy that outlines how production companies should deal with language issues in programming. S4C's language policy, or its *Canllawiau Iaith* ('language guidelines') (S4C, 2008)⁹¹, form part of its compliance guidelines⁹². Compliance guidelines are written for S4C suppliers to ensure that the content produced for S4C is acceptable to S4C. They relate, for example, to offensive language, stereotyping, political (im)partiality, etc. Language, then, is highlighted through formal institutional policy as having the potential to affect how S4C is perceived.

The guidelines reveal tensions similar to the ones elucidated above – tensions between the (purist) standard language ideology and serving a (meta)linguistically diverse viewer-base whilst maintaining the values of pluralism and inclusivity. Amongst concerns for clear articulation and simple language (attending to accessibility), the language guidelines also talk about language in more evaluative terms. They mostly allude to notions of correctness and standards, but other evaluative concepts are used, e.g. "ieithwedd...(g)ywir a(c)...naturiol" (§1.5, 'correct and natural idiom'), "iaith sy'n safonol, yn gywir" (§1.6, 'language that is standard, correct'), "geirfa Gymraeg gyfoethog" (§2.3, 'rich Welsh vocabulary'). However, the guidelines offer very little guidance on what is correct/standard and what is not. With the notable exception of the negative evaluation of contact effects, the only further definition of correctness is not particularly definitive – the guidelines ask for 'correct mutation and correct use

⁹¹ During the period the data was broadcast, S4C's language policy seems to have gone through a number of revisions, resulting in a somewhat complicated set of formal language policy principles to draw on. The earliest policy, or guidelines, seem to be those embedded within the English-language version of the broader *Compliance Guidelines* (S4C, n.d. [c. 2001]). However, the Welsh-language version of this document (*Canllawiau Cydymffurfiaeth*) contains a very different set of language guidelines, which are much closer to the latest revision, *Canllawiau Iaith* (S4C, 2008) ('language guidelines', available in Welsh only). Since the 2008 guidelines cover most of the period when the data was produced/broadcast, I will use this version as the main policy resource. S4C's language policy documents are available on S4C's website:

http://www.s4c.co.uk/production/c_guidelines.shtml, last accessed 07/04/2013.

⁹² S4C also has a Welsh Language Scheme, in accordance with the Welsh Language Act 1993. However, the scheme does not apply to S4C's broadcast content, and won't be discussed in relation to the data.

of noun gender' (§1.7, all translations from the guidelines are my own). No codification texts are mentioned. There is an assumption that what is meant by 'correct' and 'standard' Welsh will be understood. It is difficult, then, to conclude whether it is the language of the codification texts that is construed as legitimate. But the fact that the guidelines invoke 'good' and 'bad' language, which doesn't seem to need defining, suggests that there is an orientation to a legitimate language at work within the institution⁹³.

However, the guidelines do make some room for those who are not competent in S4C's legitimate language, as part of a pluralist and inclusive agenda. The guidelines stress that S4C wishes to be as representative as possible of the inhabitants of Wales (§1.8), including those who profess 'my Welsh isn't good enough'. The guidelines encourage producers to allow members of the public of *all* language competencies to contribute to its programming. However, members of the public are not treated the same as 'professional contributors', who are most central to the self-presentation of S4C (they are called '(Main) Faces' in S4C's updated Programme Guidelines (S4C, 2010)). Professional contributors are expected to use a 'high standard of correctness and articulation', including correct mutations and noun gender, and to have a 'good grasp of Welsh idioms, being able to avoid needless literal translations from English' (§1.7). So speakers whose language doesn't conform to what is considered correct within S4C are only represented at the margins of S4C's broadcast voice.

After a general first section, the guidelines move on to deal with language contact specifically, in particular in relation to vocabulary. Indeed, they dedicate an entire section to 'The Occasional Use of English Words or Clauses'. Here again the guidelines are suggestive of a legitimate language at work within the institution, since they construe English words and phrases (lexical borrowing or code-

⁹³ This is in contrast to the earliest version of the guidelines. The English version of the *Compliance Guidelines* (S4C, n.d. [c. 2001]) makes a clear statement that "It is not S4C's responsibility to standardise language nor to impose a single linguistic norm" (§5.3). This statement has been removed from later versions. Neither is there any reference in the earlier English version to notions of linguistic correctness.

switching) as incorrect, warning against their unheeded use (§2.7). And they value ‘rich’ Welsh vocabulary, even seeing for S4C an educational role in enabling access to new Welsh vocabulary (§1.10). This section is also very unclear on the definition of ‘rich’ Welsh vocabulary, other than its not being English. Even established (in speech) borrowings are discouraged (§2.1). This section doesn’t refer to any codification texts dealing with vocabulary. It is not clear, then, to what extent the presumed legitimate language is to be found in recent codification texts of lexical planners. On the other hand, there is an acknowledgement of the subjectivity of ‘what is considered acceptable, desirable or commonplace’ (§2.4) in terms of vocabulary. In this section, the guidelines grant that different people will have different competencies *as well as* different metalinguistic conceptualisations, a concession that isn’t made in the previous section. It is suggestive of an awareness that legitimacy is not black and white.

Finally, this section again seeks to balance the need for purism and correctness with the need to ‘ensure that the language used in a programme is completely accessible to its target audience and assists that audience in responding positively to and in enjoying the programme’ (§2.3). The previous section has already given some advice on how to ensure that vocabulary is accessible, in particular when ‘unfamiliar or new technical terms’ are used: by ‘avoiding a series of unfamiliar words in succession’; ‘avoiding the use of unfamiliar words alongside constructions that don’t belong to the spoken language’; and ‘seeking to consolidate understanding of new words by ensuring that the image or context supports the meaning’ (§1.10).

That such a document exists suggests that S4C sees language as fundamental to the way it self-presents, and that it recognises that language in Wales is a sensitive issue that needs careful treatment. However, the existence of a formal policy document isn’t commensurate with implementation (Spolsky, 2004: 11), and it is not clear to what extent the guidelines are implemented. It is the responsibility of producers to ensure that programmes comply with the

guidelines. What is most likely is that the formal language policy will be brought directly into play if S4C considers that programming has gone against the guidelines to an extent that needs to be noted⁹⁴. However, formal language policies also serve another purpose. They do not merely function as a medium for laying out the rules of engagement for the information of those who are expected to follow suit. Formal policies that are made publicly available work as declarations of corporate identity. The guidelines can be seen as a statement of what S4C thinks it should be seen to be saying.

4.6 THE WORKPLACE CONTEXT

As a key site of daily interaction, or site of social reproduction in Williams' (G. Williams, 1987, 1992) terms, the workplace is increasingly becoming the focus of language planning. The 1993 Welsh Language Act required public bodies to provide some level of Welsh-language service to clients, which has led to an increasing demand for Welsh-language skills in the workplace, albeit without necessarily shifting workplace language (Hodges, 2009). Government-led language planning, however, has taken a promotion or encouragement approach to increasing the use of Welsh in the workplace. For example, the Welsh Language Board's (2009a) *Promoting and Facilitating Bilingual Workplaces* document provides guidance for public bodies on how to encourage and enable staff to use Welsh in internal communication. *Iaith Fyw: Iaith Byw* names the workplace as one of its six 'strategic areas' and commits to "Make operational standards, which will enable the Commissioner to impose duties on organisations to promote the use of Welsh in the workplace" (Welsh Government, 2012b: 39) (as well as to continue with the promotion approach). It remains to be seen what level of enforcement this will entail.

The Housing Association, where data for the workplace study was collected, operates with a language policy that prioritises the use of Welsh as the language

⁹⁴ S4C commissions independent organisations to carry out compliance monitoring, which is supposed to alert it to any infringement of its compliance guidelines.

of internal operation (as stated in its language scheme)⁹⁵. It is a policy that takes it beyond the legal requirement for the provision of a bilingual service, but one that is in tune with broader planning efforts. In this the Housing Association is fairly unique. One of the participants claimed that it is the only organisation in the area to operate through the medium of Welsh (although, as I note below, the Council claims to do so, and the Welsh Language Board (2009a) lists other organisations too). The same participant also emphasised that it is a *bilingual* workplace (that is, not exclusively Welsh-language), a statement that emphasised that all staff are fluent in both Welsh and English (a requirement for all staff, as stated in its recruitment policy), and that in its dealings with its tenants and the wider public, the Association provides a service in both languages.

No doubt the Association's policy of working through the medium of Welsh is motivated and facilitated by the fact that it is located in north-west Wales, an area which is often described as one of the 'heartland' areas of the Welsh language. 'Heartland' areas can be defined demographically – that is, those areas where the proportion of Welsh speakers in the population is highest (Gwynedd in north-west Wales was highest in Wales in the 2011 Census, with 65% Welsh speakers. Anglesey, also in north-west Wales, was 2nd highest with 57% (Welsh Government, 2012a)). Yet 'heartland' areas continue to see a decline in the proportions of Welsh speakers, as evidenced by the 2011 Census (Welsh speakers dropped by 3.6% in Gwynedd and by 2.9% in Anglesey between 2001 and 2011 (Welsh Government, 2012a)). Heartland areas can also be perceptually and ideologically defined. C. Williams (cited in Pryce & Williams, 1988) found a core perceptual culture region centring on Gwynedd, as judged by respondents from across Wales. Despite the drop in the number and proportion of Welsh speakers in north-west Wales, it is not likely that the area has lost its status as a core perceptual culture region of the Welsh language.

⁹⁵ I do not cite the Association's documentation or online sources or quote directly from these sources, in an effort to preserve anonymity.

Coupled with increasing linguistic change in the area, there are considerable linguistic tensions. In his qualitative study of language attitudes in Caernarfon, a town which has the accolade of being the last remaining Welsh-speaking town, E. Williams (2009) argues that, beyond historical colonial episodes that are still felt to be relevant, from early English conquest to 20th Century political struggle⁹⁶, more recently there is tension resulting from in-migration to the north-west Wales area, an area of outstanding natural beauty, but one of the poorest parts of the United Kingdom. It is believed that a free UK housing market puts local people at a disadvantage when buying property relative to incomers (often from more prosperous parts of England), and has resulted in inflated house prices, thus keeping much of the local population out of the local housing market. Housing difficulties for local people and increasing numbers of incomers is felt to be changing the linguistic fabric of communities and has led to language-related tensions. Language division is also captured in a network analysis based in Anglesey by Morris (1989). Morris argues that economic restructuring and development brought new jobs to the area, but that these jobs were largely taken by an imported (non-Welsh speaking) workforce. According to Morris, language became symbolic of economic struggle and, through an ideological process, local tensions were generalised so that English speakers as a whole were resented by local Welsh speakers. This, she argues, was exacerbated by the problems in the housing market.

In response to these tensions, pressure groups have campaigned for change in institutional policy, such as *Cymdeithas yr Iaith*'s campaign for a property act, as mentioned at the beginning of Chapter 1 (a campaign that is ongoing). Pressure group, *Cymuned* ('community'), campaigns specifically to protect Welsh-language communities, two of its main concerns being property and in-migration (E.

⁹⁶ Early conquest is still symbolised in the spectacular castle in and around which the town is situated. The castle was built in the late 13th century by the English king, Edward I, in his attempt to secure his conquest of Wales. The colony status of Wales, in the Gwynedd area in particular, was felt well into the 20th Century, when the Liverpool City Council was granted permission by the UK Parliament, in 1957, to flood the valley of Tryweryn, including the village of Capel Celyn, despite local opposition and the opposition of the majority of Welsh MPs and local government authorities (E. Williams, 2009).

Williams, 2009).

Apart from (or as a consequence of) these grassroots movements, there is considerable institutional support for the Welsh language in the area. For example, Gwynedd local authority's education policy seeks to "develop the ability of every pupil to be confident in both languages by the end of KS2"⁹⁷ (Key Stage 2 refers to ages 7-11), an aim it has been seeking to achieve since 1975 (Lindsay, 1993). Additional support is provided, where required, by means of intensive language courses. Welsh is also the Council's official language of administration (Gwynedd Council, 2010). Moreover, Wales-wide statutory requirements and planning efforts have led to some employment advantages for Welsh speakers, such as Welsh language officers in public bodies and the growth of the translation industry.

The Housing Association's policy of operating through the medium of Welsh can be seen as a further attempt at providing institutional support for the Welsh language, as part of their local development agenda. The Association's language policy shows considerable commitment to the Welsh language as a cultural construct, a commitment that is made clear in its current Corporate Plan and on its website. Both lay out the Association's broad aims and values, making statements as to its pride in the Welsh language and its wish to promote Welsh, as well as a commitment to develop communities where the Welsh language is integral. It is clear from such statements that a commitment to 'the Welsh language' forms a crucial part of the Association's corporate identity, but this is presented, in part, as tied with its commitment to the local community, as a housing association and social enterprise⁹⁸.

What kind of Welsh language is supported by the Housing Association? Whilst we

⁹⁷ www.gwynedd.gov.uk/gwy_doc.asp?cat=7072&doc=26126&Language=1&p=1&c=1

⁹⁸ Social enterprises are businesses that have primarily social objectives, and who reinvest profits into the business or community to enhance their social objectives. See <http://www.socialenterprise.org.uk/about/about-social-enterprise#what%20are%20ses>. Last accessed 19/09/2013.

can view support for the Welsh language in the area as tied in with localism, broader language planning efforts, especially in lexical planning and standardisation, are potentially in opposition to some local ways of speaking. Recall Jilg's (2003) study (above) in Blaenau Ffestiniog, a town in north-west Wales, where the use of Welsh words or borrowings was seen to be differentiated according to status group. The social network that was defined as engaging more with standard language domains, and that had typically gained higher educational qualifications, showed linguistic capacities closer to the kind of language promoted by lexical planners, that is, words that were defined as 'Welsh'. Those who did not engage with standard language domains typically responded with more borrowings. Another (small) study⁹⁹ suggests that language variation in the area is, in part, conceptualised in terms of class or status group differentiation. In Peblig, a large housing estate in Caernarfon and the ward with the highest concentration of Welsh speakers across Wales (88% according to the 2001 Census (Aitchison & Carter, 2004: 154)), *Pobol Peblig* ('the people of Peblig')¹⁰⁰ conducted a community consultation to investigate the community's relationship with Welsh. The consultation highlighted that participants perceived a difference between *Cymraeg posh* ('posh Welsh') and their own language, *Cymraeg Cofi* ('Cofi Welsh', *Cofi* referring to an inhabitant of Caernarfon). It seems, then, that language variation potentially indexes class or status-group membership in the area. Note also the connection here between class and locality. Ways of speaking that are not local (*Cofi*) are 'posh'. As a related example, in my own research into language attitudes (published in Robert, 2009), which asked participants to give language-correctness evaluations of voice stimuli, a respondent offered oral feedback at the end of the exercise to the effect of 'How's a Cofi like me supposed to tell how correct somebody's language is?' Standard language ideology seems to be implicated in perceptions

⁹⁹ Cited in an informal document given to me by *Pobol Peblig* detailing the findings. The document notes the findings were presented in April 2010. I have not been able to find a citeable reference to a publication of the consultation.

¹⁰⁰ *Pobol Peblig* is a *Communities First* grant recipient. *Communities First* is a Welsh Government grant scheme seeking to develop sustainable, communities and economies in Wales' most deprived areas.

of sociolinguistic variation and status group differentiation.

It is pertinent, then, to ask what cultural construct is the focus of institutional support for Welsh by the Housing Association. Are 'standard', or posh, ways of speaking given precedence, or is there a broad interpretation of Welsh supported within the institution that encompasses various local, and other, ways of speaking? It is difficult to make inferences in this regard from the Housing Association's formal policy documents. The Association's Welsh Language Scheme states that it will offer training and support, where needed, to enable its staff to work through the medium of Welsh, as well as make available lexical codification texts and Welsh-language software. We expect, then, lexical planning to be implemented to some degree within the organisation. However, it is difficult to make any other inferences from its documentation as to the language skills it requires of its staff. For example, the Association's recruitment policy makes no detailed statement about language skills, only that an 'ability to communicate' in Welsh and English are essential requirements for its staff. It does not stipulate whether an ability to read and write Welsh is necessary. Similarly, its outline of the recruitment procedure and assessment of applicants offers no guidance as to what language skills would fulfil the requirement of being able to communicate in Welsh and English. However, posts are advertised in Welsh only and documentation (e.g. application forms) are sent out in Welsh only, which might suggest to potential applicants that an ability to read and write Welsh, at least, would be a requirement.

It is worth noting, however, that the organisation runs with a hierarchical structure, with a Chief Executive Officer and Directors (together, the Senior Management Team) in higher-ranking and higher-paid posts. (There is also a Board, to whom the Senior Management Team is accountable, but board members are not drawn from the Association's staff.) There are other levels on the hierarchy, including a Management Team, down to the posts on the lowest rankings. Executive power, the power to make decisions and to speak on behalf

of the organisation, is also vested in those in higher-ranking posts. This structure is typical of contemporary business management practices. There are different roles within the Association, then. Differences in roles are not only attributed to the different kinds of jobs that are undertaken within the organisation (e.g. dealing with tenants, communications, maintenance work, finance, etc.), but to differentiated status. The question regarding language practices is whether language is implicated in the organisational hierarchy. Is there a particular kind of Welsh that is highly valued within the institution? And does this language fit within the scope of what lexical planners would promote? This is not a question that can be answered by looking at policy documents, but at how language is used within the Association.

CHAPTER FIVE:

Mediating medicine in Dr Ann's *Syrjeri*

5.1 INTRODUCTION: THE *SYRJERI* DATA

In this first section, I seek to give an impression of the data, focusing specifically on the structure and content of the meetings, and the discourse goals and priorities. My objectives are to give the reader a better idea of the nature of the data, considering what topics are covered, how to define the specialist nature of talk, and what the production company seems to be trying to achieve in the texts. There is a sample transcript, one of the *Syrjeri* items in full, in Appendix 2.

Structure and content of Syrjeri items

Each of the *Syrjeri* texts used in the data is between 6 and 8 minutes long. They all involve the health expert, Dr Ann, and a presenter (sometimes two)¹⁰¹. Eleven of the items (out of a total of twenty) also involve guests¹⁰², whose role is to talk about their experience of living with the particular condition discussed during the day's *Syrjeri*, or of caring for somebody with the condition. The items are invariably based in the studio, set up as conversations between the participants, although led by the presenters, primarily focusing on the health issue(s) under discussion for the day. I have listed the health issues discussed in the previous chapter, but they include, for example, diabetes, children's health and swine flu.

In all texts but four the same set is used. Participants sit on two sofas that are at right angles to each other. Behind the sofas is a screen which shows the *Wedi3* logo and occasionally some images that the participants use as visual aids. In front of the sofas there is a rug and a table. The background is decorated with ornate bouquets of flowers and soft, coloured lighting. The overall visual

¹⁰¹ The presenters are EJ, EW, JH and RO.

¹⁰² I do not refer to guests by their real names. Instead I refer to 'the guest in *Meningitis*', for example.

impression is informal yet deliberately styled. The remaining four texts are very similar, but with slightly different furniture.

Each item follows a similar structure. The presenter(s) introduce(s) the topic and the other participants (Dr Ann and any guests). They then ask Dr Ann questions, inquiring about causes and effects of particular conditions, what treatments are available, how common particular conditions are, etc. Dr Ann then gives a scientific-medical explanation of the condition, often using props (on-screen visuals, such as graphs, or physical props, such as a model of the human foot). Her explanations outline some or all of the following: causes, effects, diagnosis, prevention and cure. Dr Ann's explanations are largely monologues, apart from the leading questions and some questions for clarification by presenters. When guests are present, the presenters then ask them to talk about their experiences of the condition in question. Sequencing sometimes differs, with some going back and forth between Dr Ann and guests. In one case, in *Epilepsy*, a pre-recorded clip is shown on screen during the item, which involves filmed visuals and a voiceover. There are some exceptions to this format, such as a slightly different guest role, or, in the case of *Swine Flu (1)* and *(2)*, the items function as an advice forum, with Dr Ann responding to viewers' questions (questions read out by presenters).

The following gives an idea of the type of topics covered in the texts (not necessarily in this order). It is based on one of the texts, *Diabetes*, which is fairly typical of the dataset.

- i. Defining conditions, either categorically (e.g. type-one diabetes is an autoimmune illness) or, more typically, descriptively. Descriptive definitions deal with the causes and effects of the condition, usually body-internal (e.g. genetics, pancreatic failure), sometimes at a microscopic level, but also body-external (e.g. environmental factors). Effects can also relate to personal/human experience, such as pain,

thirst, weight loss, and intensity of observable symptoms. Often, the description are accompanied by graphics and props, such as a diagram of the internal abdomen or a model of the human foot.

- ii. Public health statistics (e.g. characteristics of sufferers, such as male/female, region, age, etc.).
- iii. Diagnosis and experience of healthcare service (e.g. blood tests, GP check-ups, diabetes nurse).
- iv. Treatment (e.g. tablets (including specific types), insulin, management, tests).
- v. Side-effects or complications of condition/treatment, either in terms of the inner workings of the body (e.g. blood sugar damages the blood vessels) or of outer symptoms (e.g. walking difficulties or loss of consciousness).
- vi. (Possible) related conditions (e.g. rubella, coxsackie virus). These are usually not defined in much detail.

Although there is some variation in the amount texts focus on these issues (with some, for example, focusing more on human experience), the topics covered above are most typical of the data. Typically, Dr Ann does the work outlined, in particular the body-internal descriptions of the conditions. However, presenters do some defining work, in particular in their introductions, and often give some public health statistics. Guests typically share with Dr Ann the work of describing the human experience perceptions of the condition (what is observable to them, such as pain or tiredness), as well as experience of healthcare. However, guests may give some more specialist definitions also, since they have some specialised knowledge of the conditions, based on their personal experiences of the conditions and related healthcare services.

Given the range of topics covered, and the range of concepts that can be considered 'health-related' in some way, some vocabulary is more obviously specialist than others. For example, the *Bunions and Corns* text includes a number

of words that are clearly health-related and central to Dr Ann's definitions of bunions and corns, but that are very familiar as part of the 'voice of the lifeworld' (Mishler, 1984), what Fairclough (1995b: 10) explains as "a commonsense world of ordinary experience": *troed* ('foot'), *bysedd traed* ('toes'), *coes* ('leg'), *asgwrn* ('bone'), etc. Similarly, there are words that relate to health in that they convey the human experience of (ill)health and medical treatment, such as *poen* ('pain'), *lleddfu* ('to alleviate'), etc. On the other hand, there are vocabulary items that seem to be more specialist in nature, since their definition, and their designation, are embedded within medical/scientific discourse. For example, *myelin* (a type of fat found around the nerves) is specialist since 'myelin' is only known as a result of scientific/medical research methods. It is an object of knowledge (Goodwin, 1994) in scientific/medical discourse. In a similar vein, *artery*, in its meaning of 'blood vessel that carries blood and oxygen away from the heart' (as opposed to *vein*), is highly specialist. Its definition derives from medical experimentation, observation (through dissection) and theorising on the role of the heart, blood and blood vessels in the human body, the most significant part of this work having been carried out in the 17th Century (Porter, 1997: 211-6). On the other hand, even these seemingly specialist words aren't totally outside of the 'commonsense world of ordinary experience', partly because of the ubiquity of these concepts in our culture. They are objects of medical knowledge in some ways, yet known in the ordinary world. The word *artery* is very familiar, even if its specific, medical definition is not always understood or invoked by non-experts. Moreover, compulsory primary and secondary education means that the understanding of the specific function of arteries is not restricted to medical experts, even if this understanding originally derived from medical research. *Artery* can be known even in the ordinary world in the same (or very similar) way that it is known to a medical expert.

These issues highlight the potential difficulty in the differentiation between the 'lifeworld' and medical discourse, a point also emphasised by Sarangi & Clarke (2002). The practice of medicine is rooted in the lifeworld and, in turn, affects it,

and it is not always clear where lifeworld ends and medicine begins, perhaps because they do not inhabit distinctly separate space. Specialisation is not a concept that can easily be pinned down, and the distinction between medical world and life world is increasingly difficult to make. Nonetheless, these can be useful categories, provided that their subjectivity is born in mind. We find a range of vocabulary items in the data that are related to health in some way, but that can be difficult to define as 'specialist' or not. What is 'specialist' and what is 'ordinary' is a subjective matter. As such, it is likely something that participants have to negotiate in interaction, to the extent that this is possible with an audience that is presumably diverse and is not co-present.

Discourse goals and priorities

Syrjeri items seem to have three priorities common to all items: 1) to give expert information on matters relating to health; 2) to convey this information in a manner that makes it accessible to non-experts; and 3) to entertain. Together, these purposes require the merging of medical discourse with the lifeworld, or commonsense ordinary experience. However, the entertainment value also reminds us that the backdrop to the interaction is neither medical world nor lifeworld, but the broadcast media. These are media texts, and whilst they draw on medical discourse and common, everyday experience, there are interactional and institutional priorities that are related specifically to the media environment.

The mixing of these three worlds seems to be built into the design of the texts through the strategic use of the participants – presenter, expert and guest. Dr Ann, as expert, is used to give her expert knowledge, but does so in a way that suggests that she is attending to the crossover between medical discourse and the ordinary experience of non-experts. Guests are also used to help make expert knowledge accessible to non-experts through their personal experiences, as well as to intensify the human interest element and thereby the entertainment value. For example, in *Meningitis*, we are introduced to a little girl who contracted meningitis as a baby and who survived against the odds. Her story seems to be

designed to pull on the heartstrings, and to serve as a moral that there can be a happy ending to a meningitis story (the latter is expressly stated by EJ). The presenters, on the other hand, play a number of roles, being responsible for directing talk, and ensuring that the interaction meets the requirements of the production company. Their roles include: helping Dr Ann and guests relate their expert/entertaining explanations and stories; conveying some expert knowledge (introducing the topic with a number of pre-selected facts); facilitating the lifeworld-friendly presentation of expert knowledge, by taking on the role of non-expert (e.g. seeking clarification).

The merging of the three worlds – lifeworld, media and medicine – is also evident in the different genres the texts draw on. I take *genre* to refer to “schemata or frames, embodying presuppositions associated with ideological values and principles of communicative conduct that in a way bracket the talk and thereby affect the way in which we assess or interpret what transpires in the course of the encounter.” (Gumperz, 1999: 456). The emphasis here is on bracketing talk (speech events or even sequences within events) and how that talk is interpreted according to a particular schema, and not on fixed and unitary rules of conduct. The *Syrjeri* items switch between the presentation style of Dr Ann’s expositions, complete with props and on-screen visual aids (helpful for explanation and entertainment), and the personal experience-based narratives of guests. This is woven together by the presenters’ question prompts, through which presenters control the floor. The effect is part informative presentation, part television interview and part chat between friends¹⁰³. I do not mean that this blending of genres is an innovation of *Wedi3*’s. In particular, it is a very common for television interviews to be presented as friendly chats, and this could be considered a genre in itself. But it is useful to draw the reader’s attention to the fact that a television interview needn’t be a friendly chat, that participants aren’t

¹⁰³ Whilst a television interview and a chat may seem very similar, they are distinguishable since there is usually an imbalance of power in a television interview, with presenters (as well as the production team) controlling the interaction. Consequently, television interviews can be formal affairs (although in recent times they don’t tend to be).

on an equal footing, despite appearances, as well as that a chat between friends doesn't necessarily involve question prompts and extensive, explanatory answers. All genres are apparent in the set: the presentation genre by means of the screen and its visual aids; the television interview by means of the placement of participants opposite each other (albeit not directly); and the friendly chat genre by means of its living room furniture. The latter is also embodied by the participants, who avoid sitting directly facing each other or facing the camera (other than in the initial introduction to the topic), and who have their legs forever crossed.

The 'friendliness' of this style can be seen as an example of what Fairclough (1995a, 1995b) has called *informalisation*, the bringing of private discourse into public, caused by a shift from authority-based relations to negotiated identities (Fairclough, 1995a: 137). However, the style of the show is not completely informalised, or friendly-chat-like. Whilst the set makes use of living room furniture and soft lighting, there is a distinct impression of orderliness. There is no living room clutter, and clothes, hair and make-up are highly styled, especially the presenters'. Similarly, in the interaction, turn-taking is clean. That is, there is some back-channelling, but otherwise, very few interruptions or cross-talk, and the interaction mostly follows a question-and-answer format. Moreover, there is little laughter, challenging, sarcasm or irony (one or two examples), any action that could be interpreted as contrary or uncooperative, and swearing is absent (and strictly regulated through institutional policy). The style is distinctly straight-laced. This reminds us that, although the interaction is chat-like, it takes place in the media world, a world where semiotic representation is highly salient, and closely attended to by social actors who have gained skill in this work through training and years of experience.

5.2 SUBJECT VOCABULARY AND CODIFICATION TEXTS

Deciding which lexical items to include and which not was challenging. It clearly made little sense to include all words, or even all lexical items, given that my

interest in this data was partly due to its semi-specialist focus. And yet, given that the world of medicine and the lifeworld are so intertwined, in practice it is difficult to decide what talk belongs to which world. Take, for example, a short extract, taken at random, from *Epilepsy*, spoken by Dr Ann:

“ie (.) wel mae'r (.) ymennydd (.) yn organ cymhleth tu hwnt gyda (.)
miliwne (.) o gelloedd (.) sy'n cysylltu gyda'i gilydd (.) trwy modd
trydanol”
(‘yes (.) well the (.) brain (.) is an extremely complex organ with (.)
millions (.) of cells (.) that are linked together (.) by electrical
means’)

Dr Ann's subject-matter is quite specialist here, although she presents it in a way that seeks to make it accessible to the non-expert (albeit taking for granted that the viewer has some concept of ‘brain’, ‘cell’ and ‘electricity’). Nonetheless, there are clearly health-related lexical items here – *ymennydd* (‘brain’), *organ*, *celloedd* (‘cells’). Yet there are other lexical items that don't seem to be particularly health-related, such as *cymhleth* (‘complex’), *miliwne* (‘millions’), *cysylltu* (‘connect’/‘link’), *modd* (‘means’). On the other hand, these are words that aren't out of place in medical talk, and some (e.g. *cysylltu*) are central to Dr Ann's explanation. Moreover, there is the case of *trydanol* (‘electrical’) that is clearly used here to denote a body-internal object of knowledge (electricity being the means by which the millions of cells in the brain are linked together), yet electricity it isn't typically associated with the medical world, but with other areas of scientific/technological knowledge and practice. To what extent should we consider vocabulary that is borrowed from other (scientific) subject-areas as medical or health-related? The same issue occurs with statistical talk, especially labels (more or less specific) given to categories of people for analytic purposes (e.g. *merched beichiog* (‘pregnant women’), *grwpiau o dan risg* (‘groups at risk’), etc.) and vocabulary describing (semi-)statistical trends (e.g. *mynd lawr* (‘decrease’), *cyffredin* (‘common’), etc.).

Another consideration was how to deal with lexical items that denote concepts that clearly derive from the lifeworld and yet must be considered objects of medical knowledge. These primarily include parts of the body (mostly external), such as *braich* ('arm') and *croen* ('skin'). Knowledge of these concepts might be different in the lifeworld and in medical discourse, and yet the form and denotation are the same. Moreover, in some cases words of this sort might be used to refer to the concept in a health-related way or not. For example, *braich* is used in the data to refer to an arm in a health context, e.g. in *Meningitis* when Dr Ann says "a gyda hwnna fydd poen (.) yn y breichie ac yn y coese" ('and with that there will be pain in the arms and in the legs'). But it is also used to refer to an arm in a way that is not directly related to health, e.g. in *Meningitis* where the guest is talking about when her daughter was ill with meningitis, saying "a o'n i'n (.) moyn aros yn Caerfyrddin gyda ddi yn breichie fi" ('and I (.) wanted to stay in Carmarthen with her in my arms').

There is not necessarily one solution to these questions. I decided to take a fairly broad range of vocabulary for this part of the analysis, and have sought to include all lexical items that can be considered directly related to health. I did so because, as I have argued previously, lexical planning is concerned with more than very specialist content. Lexical variation occurs amongst Welsh speakers even in the case of concepts that seem to be fairly common. For example, there is nothing particularly specialist about 'joints', although undoubtedly the concept is related to health, and knowledge of the joints can be more or less specialist. Yet we find the use of both *joints* and *cymale* in the data. Surely this is of concern to lexical planners as much as the use of both *osteoporosis* and *esgyrn brau*. In fact, G. Thomas (1991: 170) insists that more 'high profile' words are of *greater* interest to planners than more 'peripheral'. I have, then, included specialist words related to health as well as fairly common ones, such as those referring to parts of the body (e.g. *pen* ('head') and *braich* ('arm')). I have not counted their occurrence, however, unless they are used in a sense that is related to health, or as an object

of health care. Thus, I have included lexical items that refer to parts and functions of the body, including those that are anatomic and those that are microscopic, as well as words referring to experience of ill-health and maladies and their causes. I have also included lexical items referring to healthcare services, such as treatment, diagnosis, testing, and the institutional/professional labels (*consultant, rheumatologist*, etc.). I have not included words borrowed from other fields of knowledge, either fields that are closely linked to medicine (e.g. statistics), or fields brought into mind through conceptual metaphor¹⁰⁴. For ease of reference, I refer to the lexical items that I have identified as health-related as 'terms'. It should be born in mind, however, that these lexical items are labelled 'terms' for my own analytic purposes. I consider that whether or not a word is a 'term' is not an objective fact, but relates to the degree of specificity of reference and the background knowledge needed to understand the meaning of the reference, something which is not given but negotiated through interaction.

After selecting the terms for analysis, I consulted a number of reference texts to investigate the degree to which the words used in the data can be said to be 'codified' (i.e. are they in line with the recommendations of the codification texts?). Amongst these texts there were general as well as subject-specific dictionaries and glossaries. I have outlined the general codification texts in Chapter 3, but here I will say a word about the subject-specific texts. The field of health is vast, consisting of a number of specialist sub-disciplines, drawing on the scientific method for its epistemological lens, working through an intricate and complex system to provide health services for the public. No one glossary or dictionary can deal with health in its entirety. Moreover, knowledge of health and healthcare provision are in a constant state of flux. Since codification texts take time to compile, it is hard for planners to keep abreast of changes. Lexical codification texts are always historical texts to some extent. Furthermore, the

¹⁰⁴ Dr Ann makes some use of conceptual metaphor in her explanations. For example, in *Diabetes*, in order to explain the role of glucose in the body she likens the human cell to a factory, which needs energy in the form of petrol to function. Glucose, she says, is the body's petrol. Thus she uses the representational field of mechanical engineering – familiar in the lifeworld courtesy of the ubiquitous car – in order to represent knowledge of the inner workings of the body.

participants in the data discuss a number of topics related to health, meaning that they draw on a number of the specialist sub-disciplines as well as all the other facets of the field.

Beyond the general or multiple-subject dictionaries and glossaries discussed in Chapter 3, there have been numerous corpus planning initiatives seeking to standardise terms specifically in healthcare, or in fields closely related to healthcare and medicine. Much of the work to date has involved the Terminology Centre at Bangor University (see Chapters 3 and 4), who typically use what Prys has called a “‘bottom-up’, demand-led approach” (Prys, 2007b: 116) in its terminology projects. This has meant cooperating on terminology projects with organisations and practitioners within the specific fields that are the focus of terminology projects. The Centre has worked within the *North Wales Health Care Standardization Terminology Group* on a number of health-related terminology projects, such as *Terms for Health Promotion* (Prys, 2000b) and *Child and Adolescent Mental Health Terms* (Prys, 2002b). Beyond representatives from the Terminology Centre, membership of the Group includes healthcare professionals, language professionals, academic researchers and university staff with responsibility for health-related provision, and staff at public organisations with responsibility for healthcare (Prys, 2007a). By involving healthcare professionals and other representatives from health-related services, implementation measures are built into the design of the projects, enabling professionals to have “a part in formulating Welsh terms, and therefore [to] feel ownership for those terms, and an interest in using them” (Prys, 2007b: 116). Presumably, these glossaries are more likely to be publicised within the organisations involved, and thereby the terms codified more likely to be used. In terms of wider dissemination, the Group specifically targets, through marketing and networking activities, other public organisations with responsibility for health services (e.g. local health boards), and for health-related training (e.g. universities), health-related professional associations (*Y Gymdeithas Feddygol*, see note 118, this chapter), language-related professional associations (the Association of Welsh

Translators and Interpreters), other public organisations (e.g. police authorities), and local Welsh-language-specific initiatives and organisations (e.g. *papurau bro*, 'district papers') (Prys, 2007a). Moreover, since the Terminology Centre seeks to integrate its various projects to ensure maximum exposure, these glossaries have been included within the *Cysgeir* resource (see Chapter 3). I decided to include these glossaries because of their likely prominence as well as their ease of use, being collated within *Cysgeir*. I also included the other health-related glossaries included within *Cysgeir*. Note, however, that *Cysgeir* does not note the source of all entries. It is clear which entries come from which glossaries.

A number of health-related fields come under the scope of the health-related glossaries in *Cysgeir*. However, I decided to include one more glossary, *Termau Meddygol / Medical Terms* (University of Wales Press, 1986), which is the only general medical glossary available. It is an English-Welsh glossary, and having been compiled in 1986, is only available as a paper text. Its not being available in digital form brings into question its usefulness and dissemination. The scope of *Termau Meddygol* is also slightly restricted and outdated, being a fairly short and old text. Nonetheless, I have used it because it is the only general medical codification text available, and since I assumed that all the more specialised medical texts, taken together, would not cover the largest part of the field that emerges in the data texts.

In summary, in addition to the general texts listed in Chapter 3, the codification texts below were consulted for the media data¹⁰⁵. The texts marked by an asterisk are all works of the *Health Care Standardization Terminology Group*:

- *An English-Welsh Dictionary of Nursing and Midwifery Terms* (Roberts & Prys, 1995)
- *Social Work and Social Care Terms* (Prys, 2000a)

¹⁰⁵ All these texts have bilingual titles. I have omitted the Welsh titles here for ease of reading. They can be seen in the list of references at the end of the thesis.

- *Terms for Health Promotion* (Prys, 2000b)*
- *Child and Adolescent Mental Health Terms* (Prys, 2002b)*
- *Terms in the Health Care of Older People* (Prys, 2005)*
- *Terms for Occupational Therapy* (Prys & Davies, 2007)*
- *Medical Terms* (University of Wales Press, 1986)

5.3 CODIFIED LEXIS IN *SYRJERI* TALK

The sample I selected for this part of the analysis consisted of four of the data texts, namely *Gallstones*, *Diabetes*, *Lupus* and *Meningitis*, which together make up a little over 29 minutes (almost 20% of the data). In this sample I categorised 161 terms as health-related, with 471 individual instances of terms (or 161 term types and 471 tokens). Table 5.1 shows the distribution of term types and tokens according to the degree to which they can be said to be codified (as discussed in Chapter 3)¹⁰⁶. In brief, I categorised each term as codified or not, depending how it was dealt with (if it was dealt with at all) in the codification texts. In a number of cases, however, it wasn't possible to categorise a term simply as codified or not, and the final categories I used are as follows: codified, not codified, partly codified, mixed, and not applicable. There is a detailed explanation of what these categories mean in Chapter 3, although I briefly explain below, where I discuss each category separately.

Category	Types	%	Tokens	%
Codified	73	45.34	282	59.87
Not codified	26	16.15	41	8.70
Partly codified	16	9.94	35	7.43
Mixed	24	14.91	75	15.92
n/a	22	13.66	38	8.07
Total terms	161	100.00	471	100.00

Table 5.1: Types and tokens of all terms in the *Syrjeri* sample

¹⁰⁶ Appendix 4 shows all the terms found in the sample, listed according to codification category.

Considering first the term types, we see that the category with by far the highest percentage is the codified category, with 73 types, or making up 45% of all term types. That is, close to half of all term types in the data are codified unambiguously as Welsh words by all codification texts consulted that had an entry for the concept. In this category, we find most of the very common body-part words, such as *braich* ('arm'), *llygad* ('eye') and *troed* ('foot'), as well as a number of slightly more expert body-part terms, albeit still very common concepts, such as *aren* ('kidney'), *cyhyr* ('muscle') and *cymal* ('joint')¹⁰⁷. We also find a number of other terms that are fairly specialist yet widely-known (or at least the *concepts* are widely known), such as *cholesterol*, *chemotherapy*, *cylchrediad* ('circulation'), *steroid* and *uwchsain* ('ultrasound'). In the case of most of these terms, whilst the concept may be known in the ordinary and medical world, the type of knowledge might vary. For example, in the case of *llid yr ymennydd* ('meningitis'), knowledge of the disease can be medical (e.g. causes and effects) or human-experience-based, including physical symptoms, suffering caused, danger of death, etc. There are no terms in this category that denote, in my opinion, concepts that are very specialist, although, as I say, medical knowledge of these concepts may differ from ordinary knowledge.

There are relatively few clear borrowings amongst the codified terms. There are approximately 17 in all¹⁰⁸, if we don't count *cell* ('cell', pronounced /kɛɭ/), *meddyg* ('doctor') and *stumog* ('stomach')¹⁰⁹. Even the so-called clear borrowings are mostly well-integrated, either phonologically (e.g. *nyrs* ('nurse'), pronounced /nərs/ and *glwcos* ('glucose'), pronounced /'glukɔs/) or morphologically (e.g. *imiwnedd* ('immunity'), built from the root morpheme *imiwn* and the

¹⁰⁷ I postulate that these words are slightly more expert since (unlike the foot, for example) they are not easily observable, hence knowable, without the means of human technology – dissection, x-ray, visual representation, etc (although some of these concepts, such as 'liver' and 'kidney', probably come originally from knowledge of animals). Even if these terms are not particularly specialist today, historically they were fairly specialist, and this potentially impacts on the availability of Welsh-language labels for these concepts.

¹⁰⁸ I say approximately, since I consider the matter of what is a borrowing and what is not to be largely subjective.

¹⁰⁹ These items are traceable as borrowings, but *cell* and *meddyg* are derived from Latin, and *stumog* is so thoroughly integrated it hardly feels like a borrowing.

nominalising suffix *-edd*). Those borrowings that haven't been integrated haven't been because they offer very little opportunity for integration. For example, there isn't much that can be done to *stent* or *steroid* to make them sound more Welsh. Or put differently, there isn't much about them that makes them sound foreign. (I discuss this in more detail in Chapter 4). There are also some loan translations in this category, such as *uwchsain* ('ultrasound', quite literally) and *brechu* ('to vaccinate')¹¹⁰. What we do not see in this category are clearly unintegrated borrowings, which we do find in other categories.

Table 5.1 also presents the number and proportion of *tokens* found of each term in the sample. There are 282 individual occurrences of all codified terms in the sample. That is, 60% of all the tokens found are codified, which is substantially higher (15% higher) than the proportion of term *types* that are codified. This comparison relates to the frequency that terms are repeated in the sample; there is typically a relatively high number of occurrences of individual codified term types. On average, there are 3.9 individual occurrences of each term type¹¹¹. This compares with an average of 2.9 occurrences of term types from all codification categories combined¹¹². One explanation for the higher occurrence of codified term types is that words in this category tend to be central concepts in the texts. I would expect them, then, to be repeated more often, to be more salient in preparing for the item before filming, and for researchers, presenters and Dr Ann (all of whom do some preparation work – see below) to pay more attention to the designations they plan to use for these concepts, perhaps looking them up in codification texts. For example, we find *glwcos*, *gwaed* ('blood'), *inswlin* and *cell* in the codified category, all of which denote central concepts in the *Diabetes* text, and all of which are repeated fairly frequently (11 times for all but *cell*, which is

¹¹⁰ *Brech* follows the same semantic extension as the English *vaccinate*, which suggests it is a borrowing of sorts. *Vaccine* originally referred to the cow-pox virus (from Latin for 'cow'), and was later extended to refer to a number of different substances taken internally to prevent infection. Similarly, *brech* is a virus that affects the skin, but is also used specifically to refer to cow-pox (according to *GPC*). *Brech-u*, then, literally means 'to give (cow-)pox virus' to, as does *vaccinate*.

¹¹¹ Calculated by dividing the number of codified term tokens (282) by the number of codified term types (73).

¹¹² 471 (all tokens in the data) divided by 161 (all term types in the data).

repeated 7 times). Another explanation for the relatively high frequency of codified terms is that they denote concepts that tend to be repeated across texts. These are also central concepts, but where centrality is in relation to health more generally, rather than to the particular conditions under question. The words denoting these concepts include, for example, *corff* ('body', 19 occurrences), *meddyg* ('doctor', 8 occurrences), *symptom* (15 occurrences) and *ysbyty* ('hospital', 12 occurrences), all of which occur in all texts in the sample, with the exception of *meddyg*, which occurs in all but one. Note, however, that the majority of the codified terms are still relatively low-frequency terms, with two thirds of the codified terms occurring three times or less. Nonetheless, comparing the term type and term token figures gives us a somewhat different perspective on the data, where a higher proportion of the lexis used is codified, according to the definition used for the analysis.

Whether we consider term types or tokens, however, it is difficult to interpret the proportion of codified health-related terms found in the data. 45% (types) or 60% (tokens) may seem a relatively small number, considering S4C's prioritising the use of 'rich' Welsh vocabulary (if 'rich' is to be understood with reference to codification texts). On the other hand, I have also given reasons why we might not expect the production company or S4C to be too rigid in its implementation of this requirement – the tone of the programme and S4C's overt desire to be inclusive. A somewhat clearer perspective emerges, however, if we look at the other codification categories. We see that only 16% of all term types (or 26 term types) fall into the 'not codified' category, that is, where the codification texts consulted were unequivocal that the word was not Welsh. This does seem to be a relatively small number of words that are *not* 'rich'. There are a number of words in this category that are fairly specialist, such as *rheumatologist*, *autoimmune*, *erythema*. There are a number of words that refer either to health conditions, e.g. *rheumatoid arthritis*, *SLE* ('systemic lupus erythematosus'), *glandular fever*, *salwch y blaid* ('lupus'), etc., or in some way to treatment, e.g. *keyhole (surgery)*, *rehab*, *intensive care*. There are very few words that denote very common body-

parts (only *bogell* ('belly button')), although there are some that denote ailments of the skin – *rash* and *briws* (in the case of the latter, it is unclear whether the participant is referring to more than one *briw* ('cut', 'blemish') or is using the English *bruise*, or perhaps both, considering how similar they are in meaning). There are some words that denote slightly more expert, although very common, body-parts, such as *joint* and *kidney*.

Terms in the 'not codified' category are almost all easily identifiable as borrowings. In fact, out of a total of 26 terms, only four are not clearly borrowings: *bogell* ('belly button'), *briws* ('bruise' or 'cut', 'blemish', as discussed above), *salwch y blaid* ('lupus'), and *taflu nôl* ('to throw up'). With the exception of *briws*, whose meaning is unclear, all of these forms are in some ways unfamiliar to me. *Bogell* is very similar to *bogail*, which is the term preferred by the codification texts. It is probably either a slip of the tongue by Dr Ann or it is a dialect form (most likely, since she says it on two different occasions). *Salwch y blaid* (literally 'wolf illness') is rather curious, since I have not been able to find any other instance of *salwch y blaid*, either in a codification text or in an internet search. It is most likely a term coined during the research phase for the *Syrjeri* item, where the coiner has interpreted *lupus* etymologically (*lupus* being Latin for 'wolf') and given the condition a Welsh name following the model of other names for conditions, such as *ffliw'r moch* ('swine flu', or 'pig flu'). This interpretation suggests an informal lexical policy within the production company that prioritises overtly Welsh forms over codified forms (*lupus* being the form preferred by codification texts). Finally, *taflu nôl* is also curious, since it seems to be built on the same pattern as *taflu fyny*, itself a loan translation from the English *throw up*, but uses a different preposition ('back' rather than 'up'). It is only used once in the data (by Dr Ann), and is possibly a slip of the tongue (which Dr Ann does seem more prone to than the presenters). The apparently non-borrowed terms in the 'not codified' category, then, of which there are few, seem to be anomalies.

The majority of the terms that are not codified, however, are easily identified as borrowings, and the majority of these are not integrated. Moreover, most of them carry fairly overt markings of borrowedness, at least according to 'standard'/prescriptive norms. These include borrowings that include 'English' sounds (according to prescriptive norms), such as /dʒ/ in *joint*; compounds that are made up of other borrowings, such as *autoimmune* and *ultrasound*; of these, a number of compound nouns built on an English syntactic pattern (that is, modifier-head, rather than head-modifier, which is most common in Welsh, e.g. *bile duct*, *gallbladder*, *intensive care*, *glandular fever*, etc.); and derivatives, such as *consciousness* and *rheumatologist*¹¹³. There is a small number of not-codified terms, however, that are borrowings but that do not carry overt markers, including *kidney* and *sick*. There is also a small number that have been integrated, such as *suppresso* ('to suppress') and *straen* ('strain' (of a virus)). It is worth noting that a number of the borrowings in the 'not codified' category (approximately twelve) have Welsh equivalent designations in the data. These include, for example, *autoimmune* (*hunanimiwn*), *consciousness* (*ymwybyddiaeth*), *gallbladder* (*bustl*), *joint* (*cymal*), and *rash* (*brech*). Not all the Welsh equivalents can be categorised as codified. For example, *bustl* is used to refer to 'gallbladder', although codification texts give *bustl* as 'bile' or 'gall'. However, contrasted with the borrowings in the 'not codified' category, they are markedly Welsh. Twelve is quite a high proportion of the 22 borrowings in this category, suggesting that even where such words are used they are in competition with codified/Welsh forms.

Turning again to the term *tokens* in Table 5.1, we see that there is a smaller percentage of not-codified tokens in the sample than there is of types (8.7% of tokens compared with 16% of types). Again, this relates to the frequency of occurrence of each term in the data. Differently put, on average, terms in the 'not codified' category are repeated 1.6 times across the sample, which is much lower

¹¹³ Often English derivatives are 'translated' by taking the original root morpheme and changing the derivational suffix to a Welsh one, such as *rhiwmatolegydd*, which is the preferred form for 'rheumatologist' in the codification texts.

than the 3.9 average for codified terms. *Rash* is the most frequently repeated term in this category, which occurs only five times. This suggests that terms in this category are either less central to the individual texts or less generally applicable in health-related talk. It may be that since they are apparently less central to the talk, they have not been given the same attention as the more salient words, and therefore less likely to have been given consideration in the preparation of the items. However, the relative infrequency of terms in this category may also be explained by the alternative designations in use in the data for some of these terms, borrowings in particular, as I outlined above. That is, one reason that *kidney*, for example, occurs infrequently (only once in the sample) is that it is in competition with *aren*, which occurs eight times.

Whatever the reason, similar to the codified category, comparing the proportion of term types with the proportion of tokens gives us a different perspective on the data, where not codified terms are less salient if we look at the token data, at least if we take frequency as a measure of salience.

Between the codified and not codified terms, there are a number of terms that are not so easily categorised. A total of 39% of term types in the sample fall into the more ambiguous categories of 'partly codified', 'mixed' or 'not applicable'. The partly codified category (16 term types, or 10%) involves terms where the form used in the data is almost, but not quite, identical to the codified term. I have detailed what this means in Chapter 3, but as an example, *virus* is presented in the codification texts as phonologically integrated, being spelt 'f(e)irws'. Possibly, then, the authors of the codification texts consider this spelling to represent the pronunciation /'veirus/ or /'virus/¹¹⁴. Yet this spelling might not be motivated by an assumption of how the word should be pronounced, but by

¹¹⁴ In its introductory notes, *Geiriadur yr Academi* says the following on orthography, before giving detailed guidelines on the pronunciation of graphemes: "A dictionary of the written language cannot teach pronunciation accurately, but fortunately the orthography of Welsh is, with some exceptions, broadly phonemic, i.e. as a rule one letter or combination of letters (ch, ll, rh, ng) represents one phoneme. By following the rules of orthography, Welsh words may be pronounced at sight, giving a standardized pronunciation intelligible to all educated Welsh speakers" (Griffiths & Jones, 1995: xx).

the difficulty of representing /'vairəs/ (the form found in the data) in Welsh orthography in a manner that doesn't seem 'illiterate' (Bevan, 2008) (Bevan is former editor of *GPC*). 'Fairys' would be, I think, read off by most Welsh speakers as /'vairəs/, but as Bevan suggests, there are more than phonological considerations to spelling. It is a little difficult to determine, therefore, whether Dr Ann's use of /'vairəs/ is codified or not. There are other examples in the partly codified category that involve pronunciation, some also involve semantics and others involve small differences between data and codification texts in the grammatical words included in the expression (see Chapter 3 for more details).

The partly codified category is relatively small, with only 16 term types, and it is difficult to make many generalisations about its content. However, we can observe that it doesn't include very many common body-part terms, only *gwyneb* ('face'), which is given as *wyneb* in the codification texts but pronounced with an initial /g/ in every case but one in the data (an extremely common pronunciation, but one not favoured by prescriptivists). There is a small number of fairly specialist terms (*immunoglobulin* and *meningococcus* are not particularly widely known). But otherwise, terms are distinctly health-related, but quite common in the ordinary world (*genetic*, *pwyse'r gwaed* ('blood pressure'), *bustl* ('bile'), for example). There is a mixture of borrowings and native Welsh forms, although many of the native Welsh forms are potentially loan translations (*brechiad* ('vaccination', see note 127), *pwyse'r gwaed* ('blood pressure'), etc.). The proportion of tokens of partly codified terms in the sample is quite similar to the proportion of types (7% and 10% respectively). They are, therefore, neither particularly high-frequency nor low-frequency, and the average occurrence (2.2) is close to the total sample average (2.9). Note, however, that a number of the terms that I have included in the mixed category might also have been included here. These include, for example, *arteri*, *comesgor* ('to miscarry', or 'miscarriage'), *cholecystitis*, *systemic* and *wast* ('waist').

The mixed category involves terms over which codification texts disagree. For

example, *llawdriniaeth* ('(surgical) operation', literally 'hand treatment'), is accepted as Welsh by all codification texts except *Termau Meddygol*, which prefers *triniaeth lawfeddygol* (literally 'hand-medical treatment') or the phonologically integrated loanword *operasiwn*. Terms that might have also been classified as partly codified include, for example, *wast* ('waist'). It is partly codified because it is very similar, yet not necessarily identical, to *gwast*, the form preferred by some of the codification texts¹¹⁵. However, the other codification texts prefer a yet more different form: *gwasg*. This is why it is also mixed in codification. There are a number of similar cases, such as *arteri*, *cerrig y bustl* ('gallstones'), *cholecystitis*, *meningitis*, *pwyse gwaed* ('blood pressure'), etc. Thus, these two categories, partly codified and mixed, overlap to a considerable extent, although I have tended to place these terms in the mixed category.

There are 24 term types in the mixed category, making up 15% of the term types in the sample. The category includes mostly fairly common concepts, although distinctly related to health, such as *artery*, *doctor*, *meningitis*, *pancreas*, etc. There is only one term that denotes what I would interpret as a less common concept, namely *cholecystitis*. A number of the terms in this category, but by no means all, are borrowings, some phonologically integrated and some loan translations. The proportion of tokens in the mixed category is fairly similar to the proportion of term types (15% and 16% respectively), and the average number of tokens per type (3.1) is close to the sample average (2.9).

Finally, the 'not applicable' category (14% of term types) involves terms for which there is no entry in any of the codification texts consulted. That is, the concepts that the terms denote are not dealt with by any of the codification texts. Such terms include, for example, *haemophilus* and *keyhole surgery*. At 14% of all term types (or 22 term types), this category is fairly large, considering that health

¹¹⁵ *Wast*, the form used in the data, may be a mutated form of *gwast*. Yet there is some variation in the pronunciation of /g/ or not in word-initial position in a number of words, with some dialects pronouncing, for example /wast/ (as non-mutated) and others pronouncing /gwast/. Codification texts tend to prescribe, through orthographic representation, pronouncing /g/ in some words (such as *gwast*) but not in others (e.g. *wyneb* ('face')).

is one of the most covered fields in Welsh lexical planning. There are glossaries in a number of sub-fields as well as fields related to health and medicine (such as a dictionary of psychology terms (Spencer, Edwards, Prys, & Thomas, 2004)). But the fairly high proportion of term types that fall into the not applicable category suggests that lexical planning simply cannot keep track of all health-related designations. On the other hand, the vast majority of the terms in this category (20 out of 22) are multi-word items, such as *llawdriniaeth camera* ('keyhole surgery', literally 'camera operation'), *peripheral neuropathy*, and *butterfly rash*. As I explained in Chapter 3, it was difficult to decide which multi-word phrases should be considered lexical items and which not. It may be, then, that lexical planners considered a number of these phrases, but decided not to count them as lexical items, or terms.

This category includes a number of quite specialist terms, such as *butterfly rash*, *meningococcus C*, *haemophilus* and *dwythen fawr* ('large duct'). Many of these are borrowings, including non-integrated loans and loan translations. Although the category is fairly large when we consider the proportion of term types, the proportion of term tokens show a much lesser frequency of these terms in the data, at 8% (or 38 tokens). On average, then, the number of tokens per term type is low, at 1.7. Like the terms in the not codified category, these are, then, probably words that are not central to the individual texts or the general health-related content of the *Syrjerei* items. If they are very specialised, this may also account for their not having been dealt with by the codification texts.

So far, I have looked at the terms in the codification categories for the entire sample. In Table 5.2 I investigate whether there is any difference in the use of terms between participants. I present the distribution of term types according to degree of codification and participant group¹¹⁶. Participants play clearly differentiated roles in the data – presenter, expert and guest. The *Syrjerei* participants also seem to reflect categories found in S4C's language guidelines –

¹¹⁶ I do not present *token* data differentiated by participant group.

'professional contributors' and 'contributors from the public'. Dr Ann is the only participant that less easily fits into one of these categories, seeming to fall between the two. Table 5.2, then, presents the distribution of terms according to codification category for Dr Ann, presenters and guests.

	Dr Ann		Presenters		Guests	
Category	Types	%	Types	%	Types	%
Codified	62	52.99	28	60.87	21	42.86
Not codified	10	8.55	4	8.70	14	28.57
Partly cod.	15	12.82	1	2.17	1	2.04
Mixed	12	10.26	11	23.91	8	16.33
n/a	18	15.38	2	4.35	5	10.20
Total terms	117	100.00	46	100.00	49	100.00

Table 5.2: *Syrjeri* term types according to codification category for participant groups

We see that it is amongst the presenters that the percentage of codified terms is highest, at 61%, followed by Dr Ann (53%) and then the guests (43%).

Conversely, it is the guests that show the highest percentage of not codified terms (29%), a figure that is substantially higher than Dr Ann's and the presenters', both at 9%. These are the most salient pieces of information to draw out of Table 5.2. There is also a substantial amount of term types in the more ambiguous categories, but the major differences between participants are to be found in the extent they use clearly codified or not codified terms.

That lexical use varies according to participant type can be explained by the different expectations placed on participants, according to type, by the institution, as well as their access to planned language. In terms of roles, S4C's language guidelines stipulate quite clearly the obligation on 'professional contributors' (here presenters) to use 'correct' Welsh. Thus despite the semi-informalised nature of the data, it would seem unusual for presenters to freely use borrowings rather than prioritise Welsh/planned words (however 'correct'

is defined). We can also assume that presenters have been selected by the production company partly for their linguistic repertoires as well as for their on-camera interactional skills. Thus they are likely to be able to deal with what S4C might consider ‘correct’ words – both familiar to them and unfamiliar. Moreover, in fulfilling their roles, they prepare handwritten notes on the items beforehand, as I observed when I spent time at the offices. They have also been given briefing notes by researchers. Therefore, even where presenters aren’t already familiar with planned/Welsh words, they will have been introduced to them before the live broadcast if they are used in the briefing notes or if they have consulted codification texts in their own preparation work (which they will have had the time to do, unlike in completely spontaneous talk). This would seem to explain presenters’ relatively high proportion of codified terms.

By contrast, the language guidelines state quite clearly that guests should not be expected to speak in any way that is not ‘natural’ to them. In line with this policy, researchers at *Wedi3* stated that, where guests express concern about their Welsh-language competencies before appearing on the show, they are told by staff that they should feel free to speak in whatever way is natural to them, and that they shouldn’t worry about using English words. That is not to say that guests do not feel some obligation to speak a certain way, from their previous experiences of television or performance Welsh, as viewers or speakers. Moreover, language competencies may be a consideration in the *selection* of guests, as researchers at *Wedi3* told me: where two potential guests were otherwise equal in what they might offer the show, “ond ma nhw’n siaradwyr da” (‘but they’re good speakers’) would be a criterion for selection¹¹⁷. However, guests’ lower proportion of codified terms and higher proportion of not codified terms suggests that guests do not feel the same obligation to use ‘correct’ Welsh as do presenters, or are not able to do so to the same extent. Moreover, guests

¹¹⁷ This is a somewhat hypothetical situation, and it may be that, in searching for guests who have experience of the (sometimes rare) conditions in question, who live fairly close to the studio, who are willing and able to come on the show, *and* who speak Welsh, the production company cannot be too picky in its selection of guests.

have not been given the briefing notes that presenters and Dr Ann get (as researchers at *Wedi3* confirmed), therefore they aren't made familiar with the words that will likely be used (Welsh/planned or otherwise). And from their own experiences of the medical conditions in question, guests are likely to have dealt with the medical profession in relation to these conditions primarily through the medium of English, since the medical world in the UK is primarily an English-language world¹¹⁸.

Dr Ann's position is more ambiguous. She makes a regular appearance on the show, and yet she doesn't have the rights of presenters to control the floor. In practice, she is likely to have been selected for her medical expertise, but it is probable that her linguistic repertoire played a role in her selection too, as well as in her continued use by the producers. In this respect, there are some media-institutional considerations to her role. Moreover, Dr Ann has been given briefing notes before each item, like the presenters, where she will have been familiarised with any Welsh/planned words intended for use, at least for the main concepts expected to come up.

On the other hand, Dr Ann also plays her professional role as medical expert on the programme. This world impacts less on Dr Ann institutionally in the media environment, but her socialisation in this world is highly relevant for the language she has acquired to function in the medical world. As noted above, it is a typically English-language world, with professional/expertise training, knowledge and practice being developed and conducted primarily through the

¹¹⁸ Language planning in the field of healthcare and medicine largely focuses on service provision, rather than on facilitating the use of Welsh as the medium of expert practice (e.g. D. Davies, 2002; E. Davies, 1999; Misell, 2000; Roberts et al., 2004). There is very little Welsh-medium professional training for healthcare professionals. At the time the data was broadcast, I found only two university departments that refer in their websites to any kind of provision of Welsh-medium professional training for healthcare careers – at Swansea University and Bangor University. (I have heard anecdotally of some language awareness training and basic language skills.) Of course, Dr Ann was trained much earlier than this time. *Y Gymdeithas Feddygol* ('the medical society') is a forum for discussing medical issues through the medium of Welsh, but is not an integrated part of the medical profession in the United Kingdom, in the same way as the *General Medical Council*, for example, that registers and monitors doctors.

medium of English. Therefore, Dr Ann is likely to have access to specialist terms primarily through the medium of English, although as noted above, she may have come across Welsh/planned terms in her media role.

Beyond considerations of institution/authority and access, however, are Dr Ann's immediate interactional priorities, which likely partly involve her mediating role. Dr Ann needs to present medical knowledge in a way that is accessible to the audience (as is prioritised in the language guidelines). This task falls primarily on Dr Ann's shoulders, since she does most of the talking¹¹⁹. In brief, Dr Ann likely has some potentially conflicting demands as well as differentiated experiences of medical language, which help explain her middle-ranking position in terms of the use of planned lexis.

Conclusions

Although it is hard to draw very concrete generalisations from the analysis, there do seem to be some trends in the data that allow some interesting conclusions. In terms of the specialisation of terms, most of the terms found seem to be distinctly health-related and marginally specialist, but still common in the ordinary world. However, as I outlined above, specialisation is difficult to define, and is largely subjective. Nonetheless, we can generalise that the more specialist terms tend to be found in the not codified or not applicable categories. That is, either they are not dealt with by codification texts, or codified designations are not used in the data for these concepts. Similarly, borrowings also tend to be found in these categories, especially non-integrated borrowings (that is, those that carry distinct markers of foreignness). However, these are only trends, and we also see more specialist terms and borrowings in other categories, including the codified category.

Nonetheless, we can conclude that a fair amount of term types are codified (and

¹¹⁹ As a rough comparison (without accounting for line length), Dr Ann has a total of 1714 lines in the data, compared to the presenters' 904 and the guests' 459.

are primarily 'Welsh' or integrated/unmarked borrowings) and a relatively small amount not codified (and primarily non-integrated/marked borrowings), and that the difference is intensified when we consider the data from the perspective of tokens. It is interesting that, although not codified terms are found in the sample, a lot of these are in competition with Welsh-equivalent terms (if not clearly codified) in the data. We find some contact-induced lexical variation in the data, then. It is not necessarily the case that the variation in designation occurs in the talk of a single speaker or even within the same text (although it often does). Nonetheless, this fact does shed more light on lexical practice, and points to the fact that, whilst terms that are not codified might be used, it is not necessarily the case that they are always used as 'normal'.

The analysis of the data according to participant type allows us to draw out more detail from the data, and to better interpret distributional patterns. As I noted above, it is difficult to interpret whether 45%, for example, is a lot of codified terms or few. Interpretation of quantitative data, indeed of any data, is often contrastive, and it is illuminating that there is a seemingly clear differentiation between participants in the extent that they use codified and not codified terms. Although I have not conducted statistical significance tests on the data, there seems to be a substantial difference between presenters and guests in the codified and not codified categories, with presenters using 18% more codified terms than guests, and 20% fewer not codified terms. Dr Ann is somewhere in the middle (53% codified terms). These differences seem to relate to participants' role within the institution and the interaction (and in the case of Dr Ann, to her professional socialisation). The differences also possibly relate to the context of text production, since Dr Ann and presenters have time to prepare and access to related texts prepared by researchers, and, potentially, to codification texts.

The extent of terms in the more ambiguous categories, however (39% of types and 31% of tokens) throw open some questions about the nature of lexical

planning. The terms that I have called ‘partly codified’ point to the fact that it is not possible for lexical codification texts to fully stipulate how language ought to be used, at least not in spoken practice. To some extent, planners do not overtly seek to influence spoken practice, although I have speculated in Chapter 2 that to view lexical planning as a wholly written-language targeted endeavour is too simplistic. Certainly, it seems from the data that participants do orient in some way to lexical planning in their spoken practices. But my main point here is that there are always some grey areas where it is difficult to stipulate or to interpret codification texts.

The mixed category points to the possible conflicts between codification texts, where one text stipulates one form and another text rejects it in favour of another. I noted in Chapter 3 how there was a certain amount of competition between lexical planners, and the mixed category is suggestive of this competition. This relates directly to standardisation. Standardisation, as I have argued, is not a matter of merely compiling a list of terms and giving them equivalent designations in another language. It relates to the way these forms and the codification texts are presented, disseminated, and oriented to as ‘correct’ or ‘legitimate’ within certain (imagined) bounded contexts (such as a register or a national language). Thus, inconsistency between codification texts causes potential confusion as to what is ‘correct’ and what isn’t, something lexical planners seek to avoid. Of course, participants themselves may not be aware that there are different recommendations in different texts; there may be no confusion as far as participants are concerned. A degree of uncertainty has inevitably been built into the design, by my decision to consult a number of codification texts as well as codification texts of varying natures (terminological and lexicographical). My decision to include these various texts was based partly on the fact that no one codification text was likely to cover all the terms that came up, but also on the fact that I felt that no one text could be considered wholly authoritative.

Some of the data also points to the apparent covert/informal policy within the context of the production company of prioritising Welsh forms over codified. We see this, for example, in the case of *salwch y blaid* (Welsh, not codified) and *lupus/lwpws* (borrowing, codified). We also see it in the case of Welsh words that *are* codified, but where codification texts stipulate a meaning different from that used in the data, such as *bustl* ('bile' according to the codification texts, but used for 'gallbladder') and *brechiad* ('vaccination', used for both 'vaccination' and 'vaccine'). This allows us to speculate that participants don't always consult codification texts, but use their own intuition and linguistic resources to decide what words to use and what is 'Welsh'. Alternatively, we might infer that they do consult codification texts but find them difficult to interpret.

Finally, the not applicable category, the terms that are not dealt with by any of the codification texts, point to the fact that codification texts don't deal with all designations within a certain subject-area. This is partly because of the sheer extent of lexical designations in health-related fields, and because of the indeterminacy of registers/domains. Register and domain are abstracted concepts, and seeking to codify language within these constructs necessitates drawing a conceptual line around language where none exists, objectively speaking. The not applicable category also reminds us of the fluidity of language. Codification seeks to fix language; 'for all people for all time' as Cooper puts it (1989: 131), drawing a (limited) comparison with the standardisation of weights and measures. But the sheer number of possible concepts and designations (which are, arguably, infinite) – including compounds – means that language cannot be fixed. Codification attempts can, at best, be partial.

5.4 IN SITU LANGUAGE PRACTICE AT THE SYRJERI

The above analysis has given some abstracted impressions of the distribution of terms across the data, and between participants. What the discussion suggests, with its many nuances and generalisations, is that a qualitative understanding of the data would be valuable. In some ways, the above analysis is qualitative;

despite being an interpretation of some quantitatively presented data, interpretation has also largely involved a discussion of the words within each category in relation to specialisation and borrowing. However, whilst a quantitatively extracted, decontextualised, set of secondary data offers one perspective on the use of lexis, this type of analysis leaves some issues untouched or in need of further examination. For example, the above analysis has allowed me to identify that a number of the non-integrated and not codified borrowings coexist in the data with more 'Welsh' terms – lexical variation. However, the analysis so far doesn't explain *why* this happens, or what the variable lexical forms might mean within the context of the interaction. This brings me back to my research questions: whether planned lexis is used, what social meanings seem to be activated in the use of planned (and unplanned) words, and what local contingencies constrain (or enable) their deployment.

I argued in Chapter 3 that an analysis of social meaning in relation to standardisation required an approach to normalisation. I postulated that normalisation could be investigated through Fairclough's (1995a) notions of naturalisation and orderliness. What language code is presented as (or allowed to be) naturalised in this data through orderly interaction? I primarily consider orderliness by looking at conversational structure (whether interaction flows according to institutional norms, whether what is said is contested, etc.). However, I begin by considering the role of key in relation to orderliness. Keying is a notion that Coupland (2007) borrows from Hymes' (1972) SPEAKING mnemonic, to refer to the manner in which something is said, for example serious or ironic, which allows us to infer the speaker's communicative motivation. Key is perhaps most salient where the effect of what is said is markedly different from the propositional meaning of the utterance, such as in the case of sarcasm. Key is relevant to orderliness because a certain key can subvert the discourse – challenge the dominant ideology – without the challenge being explicitly uttered through lexicogrammatical resources. We must be sensitive to key, then, in making inferences about orderliness.

Less salient is when keying aligns the effect of what is said with the overt content of talk. But it is clear that in the vast majority of *Syrjeri* talk the key is straight-down-the-line. As I noted in section 5.1, *Syrjeri* items are straight-laced, informal yet orderly. I am tempted to call this key serious. And yet, it is not that the interaction is grave. Perhaps a better designation would be *earnest*, a concept that alludes not so much to the gravity of talk as to its apparent sincerity and to the speaker's identificational authenticity. In earnestly keyed talk, the speaker identifies directly with what is said, or projects as much (earnest keying is, of course, used by those who seek to deceive). In my analysis below, then, we can assume that there is no disorderliness projected through keying.

However, there is one very salient example of talk in the data that is not keyed earnestly. It is significant, I think, that it is the phonologically integrated borrowing, *syrjeri*, and not the 'native' semantic equivalent, *meddygfa*, that has been selected as the title for the items themselves. *Meddygfa* would not work here, probably because it would cause the key to be too serious. It would not fulfil the same function that *syrjeri* does¹²⁰. It is not that the use of *syrjeri* suggests that the speaker (or perhaps *Wedi3*) is not to be trusted. Rather, the designation seems to be tongue-in-cheek, framing the *Syrjeri* items within the wider context of the light-hearted entertainment programme. There is a methodological difficulty here, of course; it is very difficult to infer, as an analyst, the meaning of *syrjeri* in this context without drawing on a lot of my own background knowledge. However, this use of *syrjeri* seems to draw on a strong tradition in Welsh-language light-hearted entertainment, which I first encountered at a young age, primarily on S4C, where the use of English borrowings has been largely restricted to non-earnest talk. It is an observation also made by Ball et al. (1988: 193), albeit they characterise such practice as metaphorical code-switching. They list as examples: expressing irony, satire, humour, anger and

¹²⁰ Note, however, that *syrjeri* is also used in earnestly keyed talk. *Meddygfa* and *syrjeri* (with the specific meaning of 'GP's surgery') are each used twice in the data, other than in reference to the *Syrjeri* items.

disagreement. These examples could all be viewed through the lens of key.

I turn now to the bulk of the *syrjery* data, where keying is earnest, first considering a data extract in which, I argue, we see orderly interaction – interaction that proceeds according to participants’ expectations, including expectations in relation to the language code. In Extract 1, taken from *Gallstones*, Dr Ann is in the middle of one of her informative presentation-like expositions, which she is carrying out with reference to an anatomic diagram on the screen. The diagram shows labels of a number of parts of the anatomy, some of which Dr Ann uses in her talk. Dr Ann is explaining the function of the gallbladder and what happens when a gallstone blocks the passage of liquid (I assume bile) through the bile duct.

Extract 1

1. DA: nawr welwch chi mae **dwythyn fawr** yn mynd o’r **afu** (.) mewn i’r
now you see a large duct goes from the liver (.) into the
2. **berfedd** a pwrpas y **bustl** na (.) ydy casglu hylif o’r **afu** (.) sy’n cael ei
intestine and the purpose of that gallbladder (.) is to collect liquid from the liver (.) that is
3. gynhyrchu (.) ar gyfer trin y bwyd //(.) **braster**\ pan chi’n cael pryd
produced (.) for treating the food (.) fat when you have a meal
4. EJ: /reit (.) ie\\
right (.) yes
5. DA: o fwyd yn enwedig **braster** (.) mae’r **bustl** na yn gwasgu (.) ac yn (.)
especially fat (.) that gallbladder presses (.) and (.)
6. arwain (.) hylif (.) mewn (.) //nawr te\
leads (.) liquid (.) in (.) now then
7. EJ: /trwy’r goeden\\ uh //werdd na\
through that uh green tree?
8. (EJ is referring to the depicted gallbladder and the ducts that lead from it to the liver and stomach, which are coloured green)
9. DA: /trwy’r goe\\den werdd na (.) nawr welwch chi a- (.) mae na (.)

through that green tree (.) now you see (.) there's (.)

10. **dwytthen fach** yn mynd o'r **bustl dwytthen fawr** yn dod o'r **afu** (.)
a small duct going from the gallbladder a large duct coming from the liver (.)
11. nawr os oes bloc yn y **dwytthen fawr** na (.) mae'r **afu** yn mynd i
now if there's a block in that large duct (.) the liver's going to
12. diodde a chwyddo (.) a chi mynd i gael **haint** (.) a chi'n mynd yn felyn
suffer and swell (.) and you're going to get an infection (.) and you turn yellow
13. (.) mae'ch **carthion** yn mynd yn wyn mae'ch dŵr chi'n mynd yn
(.) your stools turn white your water goes
14. dywyll (.) ch- **jaundice**
dark (.) ch- jaundice
15. EJ: mmm

Given EJ's role as a presenter, who is tasked with ensuring that the interaction proceeds according to institutional norms/priorities, and who also has the rights to control the floor, she is in a position to clarify or correct Dr Ann's use of language (to the extent that she is able to follow Dr Ann's talk). In fact, she does confirm understanding ("reit (.) ie", line 4), signalling, or cueing, that Dr Ann's talk is proceeding as normal. She also clarifies Dr Ann's meaning in line 7, by means of a metaphor, drawing a likeness between a part of the diagram on the screen (ducts, coloured green, leading into the liver, and branching out) with a tree. Here she cues that Dr Ann needs to pay closer attention to representation to a lay audience, but does the clarification work for her. Note, however, that in terms of the language code itself, EJ does no correction work here, suggesting that the interaction is proceeding as normal. It is the specificity of talk that EJ is calling attention to.

There is a certain sense in which it might be difficult for EJ to correct Dr Ann's language code, of course, since it is drawn from a profession different from EJ's. But as I have already highlighted, health-related talk is closely intertwined with

the lifeworld. Moreover, *Syrjeri* talk is largely planned in advance, and lexical content, at least that which is salient within the interaction, can be vetted beforehand, and the expected vocabulary made available to EJ (I elucidate on this below). EJ, and other presenters, come to the interaction prepared. Another factor that might make it difficult for EJ to correct talk is the face-threatening nature of other-targeted repair, in particular on the grounds of ‘correctness’ (as opposed to clarity of meaning). We will see later other examples where EJ does not correct participants. I infer that EJ’s primary role is facilitator of talk, and correcting others and threatening their face directly is more likely to interfere with the narrative than facilitate talk. Nonetheless, the point is that facilitating talk from Dr Ann and the guests is given higher priority in interaction than insisting on the use of a particular language code. Dr Ann’s talk does not impinge on EJ’s sense of appropriateness of language code to the extent that she would interfere in the narrative to correct it (as she might if Dr Ann swore, or started speaking Japanese, for example). In this way, Dr Ann’s talk is *allowed to come into* the discourse (Fairclough, 1995a; Shapiro, 1989), and is presented as normal.

The health-related lexical items in this extract (which I coded for the quantitative analysis) are: *afu* (‘liver’), *braster* (‘fat’), *bustl* (‘gallbladder’), *carthion* (‘stool’), *dwythen fach* (‘small duct’), *dwythen fawr* (‘large duct’), *haint* (‘infection’), *jaundice*, *perfedd* (‘intestine’). All these words, then, pass as normal within the interaction. In this list we see one borrowing, *jaundice*, a word that is not codified as Welsh in any of the codification texts consulted¹²¹. We have, then, quite a clear case of a borrowing that passes as normal in the *Syrjeri* texts, but that is not accepted as Welsh by lexical planners. EJ even confirms her understanding here with an interjection (line 15) during a marked pause in Dr Ann’s talk, a transition-relevance place where she could have intervened. (Dr Ann then continues with her explanation.) *Jaundice* is an example of a not codified term, a partly marked borrowing that seems to be allowed within this very orderly interaction.

¹²¹ It is, however, included in *GPC* as *jandis*.

Of the other health-related words in this extract, none are clear borrowings, but not all are accepted by codification texts. In fact, only four (*afu*, *braster*, *carthion* and *haint*) are unambiguously codified as Welsh¹²². The others fall into my ambiguous categories above. These include, for example, *bustl*, which, as we have seen, is codified but not in the meaning used in the data ('gallbladder'), but with the meaning 'bile'/'gall' (the liquid). *Dwythen fach* and *dwythen fawr* are not found in the codification texts at all. Although it is likely that these labels are simplifications¹²³, which might explain their absence from codification texts, these expressions are arguably presented by Dr Ann, within the context of the diagram labelling and the other lexical items presented, as terms relating to the liver and gallbladder. Moreover, *dwythen* by itself is not found in codification texts either, the similar form *dwythell* being preferred.

What we see in Extract 1, then, is the presentation of lexical items as normal, without these words necessarily being those preferred by codification texts. What most of these lexical items have in common, however, is that their morphological content is Welsh in origin¹²⁴. They may not all conform to the preferred norms of lexical planners, but they are in line with purist ideology. The one exception is *jaundice*. It seems, then, that institutional norms share the same purist ideological base as codification texts, but codification texts are not seen as absolute authorities. This is a point I have already suggested above (see p.147 on *salwch y blaidd*).

I emphasise that the naturalised presentation of lexical items in orderly

¹²² It is worth noting that *afu* ('liver') is typically viewed as southern. This is corroborated by codification texts, who all offer *iau* (typically viewed as northern) as an alternative.

¹²³ The duct that Dr Ann refers to as *dwythen fach* ('small duct') is given in *Gray's Basic Anatomy* (Drake, Vogl, & Mitchell, 2012: 170) as the *cystic duct*. However, it is difficult to infer whether her *dwythen fawr* ('large duct') relates directly to one of the specifically labelled ducts (possibly the *common hepatic duct*) or to the collection of ducts that carry the bile. It is possible that Dr Ann's labels here are more descriptive – pointing to the tube-like structure in the graphic that is larger, and the one that is smaller.

¹²⁴ With the possible exception of *bustl*. *GPC* suggests the word is of Celtic origin, and yet compares its form with the Latin *bilis*.

interaction in the *Syrjeri* data does not necessarily mean that viewers perceive lexical usage as naturalised, or indexing nothing. Planners, for example, were they to watch *Wedi3*, might realise that Dr Ann's (and the graphic's) use of *bustl* is 'incorrect' (from the planner's perspective). Other viewers are likely to consider the use of *jaundice* 'inappropriate'. Others still might find the use of *carthion* pretentious or high-brow, never mind its denotational meaning. Orderliness and naturalisation relate to what is allowed to come into the discourse, and not to how talk is likely to be judged by those largely outside the interaction. Of course, viewers are part of the interaction, but only in a very marginal sense – as an imagined audience. Viewers have very little direct influence on what is said, beyond how S4C might respond, for example, to letters of complaint, viewer figures, or comments made during S4C's viewers' evenings (see Chapter 4).

So far, I have considered primarily the orderliness of talk – whether lexical items are used in interaction without leading to comment or conflict. What about interaction that cannot be identified as orderly? Fairclough (1995a) suggests that *disorderliness* is a sign of competing norms within an institution. There are some possible cases of disorderliness in *Syrjeri* talk – where there seems to be some contestation over norms. In the following extract from *Swine Flu (1)*, we see the use of two different designations for a single reference. This variable usage occurs by the same speaker (Dr Ann) and even within the same utterance. Might we interpret this lexical variation as a sign of disorderliness and competing norms? In Extract 2, Dr Ann (DA) and the presenter (EJ) are discussing the safety of treatments available for swine flu. In the preceding utterance, EJ has asked Dr Ann whether medication is safe for those who have an egg allergy.

Extract 2

1. DA: ydy wel y **vaccine** fydd hwnna y **brechiad** mae'n (.) mae'r **brechiad**
yes well that will be the vaccine the vaccine it (.) the vaccine
2. yn cael ei tyfu (.) mewn wye (.) felly os oes gyda chi beth ni'n galw
is grown (.) in eggs (.) so if you've got what we call

3. anaphylaxis hynna yw (.) rywbeth difrifol (.) nid jyst rash a cosi falle
*anaphylaxis that is (.) something serious (.) not just a rash and itching
perhaps*
4. ond rywbeth hollol difrifol (.) byddech chi ddim yn cael y **vaccine** (.)
but something completely serious (.) you wouldn't get the vaccine (.)
5. ond hyd yn oed os ydych chi'n feichiog (.) chi'n (.) dal yn gallu cael y
but even if you're pregnant (.) you (.) can still get the
6. **brechiad**
vaccine

In this short extract, which is 20 seconds long, Dr Ann makes five references (all lexicalised) to a vaccine that is used to protect humans from the swine flu virus, but alternates between both designations. In line 1, Dr Ann initially makes her reference to the vaccine with the expression “y vaccine” (‘the vaccine’), in a clause that is otherwise Welsh. However, by means of an appositive, she reformulates her reference as “y brechiad”. Then, at the end of line 1, she makes another reference to ‘vaccine’, repeating *brechiad* in a new clause. Her next reference to ‘vaccine’ comes after a gap of 14 seconds (and a parenthetical remark), where she uses *vaccine* again (line 4). But 4 seconds later, at the end of another sentence, she switches back to *brechiad* (line 6).

From a certain vantage point, this switching back and forth between designations by Dr Ann might be considered disorderly. However, let us consider the conversational structure in more detail to gain a clearer view. Dr Ann’s switch from *vaccine* to *brechiad* in line 1 is akin to what Quirk et al. call ‘revision’, in their categorisation of reformulating appositives (Quirk, Greenbaum, Leech, & Svartvik, 1985) – the speaker uses an appositive in order to slightly revise the meaning, here the social meaning. It can also potentially be seen as a type of code-switching, since it juxtaposes two words that might be interpreted as belonging to different languages. Although *vaccine* is originally a borrowing from

French, it is most likely a word that is borrowed (or switched) from English¹²⁵. In contrast, *brechiad* is a Welsh word (in its surface form), made up of the root morpheme *brech* ('pox') and the nominalising suffix *-iad*. The words, then, offer 'borrowed' and 'native' meanings, respectively, that might be activated in the interaction.

The codification texts also stipulate these meanings. From a lexical planning perspective, *vaccine* is not considered a legitimate Welsh word. All the texts I consulted that had an entry for 'vaccine'¹²⁶ were in agreement that *vaccine* was not a Welsh word. It should also be emphasised, however, that whilst *brechiad* is given as a Welsh word in the codification texts, the use of the word in the data does not conform to usage stipulated by the codification texts (hence its categorisation as partly codified above). Codification texts reserve *brechiad* for *vaccination*, the act of inoculating with a vaccine. The vaccine itself, the virus used in inoculation, is given as *brechlyn*. By these definitions, *vaccine* and *brechiad* are not semantic equivalents. However, I treat them as such here because they are treated as synonyms in the data, at least in some instances¹²⁷.

As a code-switch, we can view this example (in line 1) as a case of what is variously called in the code-switching literature reformulation, reiteration, repetition or translation (although these types of switch are not limited to appositional structures). Reformulating switches can carry out a number of functions, including emphasis, clarification and attracting attention, although as Auer (1995) insists, how the researcher interprets the switch should derive from the conversational structure as well as from what Auer refers to as 'episode-

¹²⁵ There is no entry for *vaccine* (or a Welsh spelling) in *GPC*, thus I have no information on the historical borrowing of *vaccine*, although the very lack of representation in *GPC* might suggest it is a recent borrowing.

¹²⁶ Texts consulted that had an entry for 'vaccine' were *Cysgair, Y Termiadur, An English-Welsh Dictionary of Nursing and Midwifery* and *Geiriadur yr Academi*. *GPC* does not have an entry for *brechiad* or for *vaccine* (under 'f'), but it does list *brechu* ('to vaccinate' or 'inoculate').

¹²⁷ For example, in line 2 Dr Ann notes that the vaccine is grown in eggs, which only makes sense if *brechiad* refers to the virus, and not to the act of inoculation. In the other cases of *brechiad* in the data the speaker's meaning is equivocal; they may be referring to the virus or to the act of inoculation.

external' factors, such as speaker preference for one language or the other, or 'community norms' (1995: 121).

Why does Dr Ann revise her initial formulation in line 1? One potential interpretation is that her reformulation is corrective, a self-initiated repair, where she realises that she has inadvertently used a word that is dispreferred, or counternormative, and repairs the referral with a different expression. This interpretation seems to chime with Alfonzetti's interpretation of reformulating code-switches, where the speaker "aims at correcting the use of the 'wrong' code [...] which they have almost inadvertently used" (Alfonzetti, 1998: 185). Alfonzetti goes on to explain that "the switching here highlights a conflict between norms of situational appropriateness and spontaneity of linguistic usage. It therefore provides indirect cues about the speaker's individual preferences and competences and also about the sociolinguistic evaluation of the two languages" (1998: 185). It is possible that Dr Ann's reformulation here is the result of her having *inadvertently* accessed, in spontaneous speech, her more habitualised linguistic resources, but, having reassessed the indexical potential of *vaccine* within the institutional interaction, she selects a word that she considers better allows her to perform the identity that she wishes/feels compelled to perform. She continues then with her revised formulation, repeating *brechiad* almost immediately at the end of line 1.

This interpretation suggests that, for Dr Ann, *vaccine* is more easily accessed than *brechiad*. This is related to Coupland's notion of 'choice': that in style-shifting we are seldom completely free to play out the identities we choose, since "ways of speaking [are] sometimes too ingrained in us for us to be able to opt out of their identity implications" (Coupland, 2007: 83). Dr Ann's repair, if it is a repair, suggests that she is orienting to interactional norms. But her initial formulation suggests that she has a more deeply ingrained way of speaking, that has other identity implications, and that isn't entirely easy for her to opt out of in inhabiting a different identity.

The above interpretation rests on the assumption that Dr Ann's reformulation is a self-initiated repair – a correction of a formulation that she never intended to make. By this definition, we interpret Dr Ann as articulating a dominant purist ideology, albeit having some difficulty in doing so. However, if we look more closely at the data, there is nothing in the delivery to suggest that Dr Ann's reformulation is a repair. Another possible interpretation, then, is that Dr Ann's use of *both* words is intentional. The question now is not only why Dr Ann produces *brechiad* when she has already used *vaccine*, but why does she use two words rather than one? Dr Ann's consecutive use of both words in referring to 'vaccine' may have been motivated by a phenomenon I discussed in Chapter 4, namely the tension between purism and accessibility at S4C. *Brechiad* is an example of the 'rich' Welsh vocabulary that S4C expects to hear from its presenters. Even if it doesn't conform exactly to the codification texts, it certainly sounds Welsh (as well as looks it). But Dr Ann is almost certainly attending to familiarity/accessibility by also using *vaccine* in line 1. Schiffrin (2006) uses familiarity – also known as accessibility – to refer to a recipient's ability to access what the speaker is referring to, according to how familiar the referent is from the background knowledge. Speakers do not know exactly how accessible their talk is to recipients – how familiar their recipients are with what is being talked about. However, speakers must make judgements in this respect when they are speaking – *assumed familiarity* (Schiffrin, 2006), and modify their talk to what they think their recipients will be familiar with.

Whilst accessibility or familiarity is usually conceived of as a matter of information given in the prior-text or context, we can also conceive of familiarity with reference to linguistic form – whether the recipient is (or assumed to be) familiar with the lexical item(s), in this case, used in the referring expression. This interpretation is supported by information gathered by production staff at *Wedi3*. Researchers help in the preparation for *Wedi3* items, by writing briefs summarising some of the most salient information to be included in the

broadcasts. I was told by the researcher who works on the *Syrjeri* items that they seek to prepare content 'from the perspective of the viewer who doesn't know anything'. Some Welsh words are deemed 'too complicated' or "astrus" ('difficult' or 'unfamiliar'). In such cases, Welsh and English are sometimes used side by side in the briefs¹²⁸. In particular in the case of the *Syrjeri* items, I was told, they are eager that viewers understand the content. However, it ought also to be emphasised that the English lexical items are not used instead of Welsh items that are deemed to be 'astrus', but side-by-side with them. Welsh lexical items, then, the 'rich' Welsh vocabulary in S4C's language guidelines, are given space despite their opacity.

In her use of *vaccine* in line 1, then, we can infer that Dr Ann is making a judgement that *brechiad*, the 'rich', 'Welsh' vocabulary item, is not universally accessible/familiar to the imagined audience, and therefore also uses *vaccine*, the (presumably) more familiar word. However, this conception of accessibility is ideological, since there is no objective measure that is used to determine what is and is not accessible to the audience. Language use is not a reflection of what the audience *does* understand. Rather, judgements are made about the imagined nature of the audience. Moreover, judgements of accessibility contribute to constituting the relationship between Dr Ann/*Wedi3*/S4C and the audience. By seeking to accommodate the audience, taking on much of the communicative burden herself, Dr Ann is constructing a conversation between equals, rather than positioning herself, and S4C, as an authoritative voice.

I have offered two interpretations of Dr Ann's reformulation in line 1. The interpretations differ basically in the meaning I infer from her use of *vaccine* – whether Dr Ann uses *vaccine* inadvertently, because it comes more readily to her, or because she is attending to assumed recipient familiarity. I interpret her use of *brechiad* as attending to S4C's language policy and expectations of her as a semi-

¹²⁸ Another strategy that was mentioned was the 'simplification' of Welsh, although it is not exactly clear what simplification entails.

presenter. The first interpretation – that Dr Ann’s use of *vaccine* is a ‘mistake’ brought about by relative lack of fluency in institutional norms – is plausible, in particular in light of similar examples that are evidenced in the data (which I will discuss below). However, I tend to favour the second interpretation in this case, because of the fluency of her delivery – there is nothing other than the formulation itself that suggests repair. Moreover, there are many other similar cases of reformulating appositions¹²⁹. In the vast majority of these, the word order is switched, with the Welsh origin word being presented first and then reformulated with a borrowing, such as in the following example from *Epilepsy* (a voiceover is giving instructions on what to do if somebody is having an epileptic fit¹³⁰):

“rhowch y claf yn y **safle adfer** (.) y **recovery position**”

(‘put the patient in the recovery position (.) the recovery position’)

In this example, taken from a pre-recorded (scripted and edited) clip, repair would suggest that it is the Welsh origin word that is in need of repair. This is difficult to conceive within the context of S4C, as well as considering the scripted and edited nature of the text. The use of reformulating appositions is also a strategy that has been noted by Ball et al. (1988: 193). In factual programming, where presenters prioritise what Ball et al. call a ‘high’ variety of Welsh, they tend to gloss Welsh words that are “likely to be unfamiliar to the audience [...] for example, in programmes which take the language into domains usually exclusively English” with English semantic equivalents. Nonetheless, a gloss usually comes second in the sequence (like in the *recovery position* example), not first, as in the case of *vaccine*, which leaves the interpretation of this particular reformulating apposition open.

¹²⁹ I found 57 reformulating appositions in all, the majority of which (41) involve words of different language origin.

¹³⁰ The accompanying visuals show a demonstration (not a real life incident), where a man is lying on the floor apparently unconscious, and another is treating him, carrying out the instructions that the voiceover is giving.

Before I leave the 'vaccine' extract, I also want to consider Dr Ann's further references to 'vaccine'. The reason that I have included the remaining 5 lines in Extract 1 is to show how, after initially introducing both words, Dr Ann doesn't then consistently use one, but goes back and forth from one to the other. Dr Ann does not merely use *vaccine* as a gloss for *brechiad*. She actively uses both words in making her references, as we see later in the extract where she returns to using *vaccine* in line 4, but switches back to *brechiad* again in line 6. It is difficult to interpret her lexical choices here as doing different social meaning work, at least in as far as Dr Ann's own intent is concerned. It could be that between lines 2 and 4 Dr Ann has 'forgotten' to use *brechiad* and uses the more habitualised form in the spontaneity of talk. In the next reference, it may be that reflexivity returns and she uses *brechiad* again. This would be the problem of choice and habitualised language again. Alternatively, it could be that having established both designations in line 1, Dr Ann feels free to use either word. Even where the meaning of the Welsh word *brechiad* has already been established, its use is further alternated with *vaccine* in the rest of the extract. Thus, *vaccine* functions as more than a gloss, but as equally valid with *brechiad*, as long as both are used.

To return to the question of the normalisation of planned language, if we accept that Dr Ann's use of both words is ideologically determined this would suggest that there are two (or more) competing ideologies within the institution, neither of which is wholly dominant. Whilst S4C seems to be seeking to use 'Welsh' words, as we saw in the language guidelines in Chapter 4 it does not feel freely able to do so. This is not primarily caused by the (meta)linguistic diversity of its audience, however, but by S4C's perceived *need to accommodate* tracts of targeted potential viewers who might not usually engage with S4C. Purism, then, does not seem to be wholly naturalised in the *Syrjeri* data. Moreover, as in the case of a number of the words found in Extract 1 above – *jaundice*, *perfedd*, *dwytthen fawr*, etc. – *brechiad* is not the preferred designation of lexical planners for the concept 'vaccine'. It is not, then, the norms of lexical planners that are normalised (in part) in this extract, but purist ideology.

Whatever the motive behind Dr Ann's reformulating apposition in Extract 2, two themes emerge from the interpretation that suggest that purism isn't wholly naturalised in the *Syrjeri* interaction – either because of the assumed metalinguistic variation of S4C's audience (and S4C's wish/need to accommodate this audience) or because of the linguistic competencies or fluency of the participants. I want to focus for the time being on the cases where Dr Ann and sometimes the presenters seem to have difficulty inhabiting linguistically the identities required for their roles within the institutional interaction. The examples below show some potential disfluency. I infer that there are three types of possible disfluency in these examples: cases of what seem to be mispronunciation (examples 1 and 2), cases where the participant seems to have difficulty finding a referring expression (examples 3 and 4), and cases where I infer that the reference is made in an incongruous way (examples 5 and 6).

1. EJ: lupus neu (.) salwch y blaidd (.) dyna'r cyflwr ry'n ni am ei drafod
 yn y syrjeri heddi tostrwydd hunanimiwn (.) lle mae'r corff yn creu
 gwrthryff sydd (.) yn lle diogelu'r corff (.) yn hytrach (.) yn ymosod arno
 fe

(‘lupus or (.) lupus (.) that's the condition we're going to discuss in the surgery today an autoimmune illness (.) where the body creates antibodies that (.) instead of protecting the body (.) rather (.) attack it’)

EJ's pronunciation is /'guθrif/ whereas she seems to be trying to say *gwrthgyrff* (‘antibodies’) (this is confirmed later, where Dr Ann uses *gwrthgyrff*). The word is usually pronounced /'gurθgirf/ (being a compound made up of *gwrth* (‘anti’/‘against’) and the plural of *corff* (‘body’)). She has possibly confused the word with *gwrthrych* (‘object’), usually pronounced /'gurθrix/ or /'guθrix/.

2. DA: a mae na gallu fod probleme gyda **biochiogrwydd**
 (‘and there can be problems with pregnancy’)

Dr Ann's pronunciation is /biɔχ-/ where usually the pronunciation is /beiχ-/

(spelled *beichiogrwydd*, from *beichiog* ('pregnant') and *baich* ('burden')¹³¹).

3. DA: mae'r llun nesa yn dangos beth mae'r llawfeddyg yn gweld (.) pan mae fe'n neud (.) **y (.) camera ma**
(‘the next picture shows what the surgeon sees (.) when he’s doing (.) this (.) camera’)

In her previous utterance, Dr Ann has presented the concept of ‘keyhole surgery’, with the reference(s) “llawdriniaeth camera beth nhw’n galw keyhole surgery” (‘camera surgery what they call keyhole surgery’). In this example she is again searching for the right label, evidenced by the hesitation, as well as by EJ’s offering her own completion with “driniaeth” (‘treatment’), overlapping with Dr Ann’s “camera ma”. Dr Ann’s use of “camera ma” is strange, since you cannot ‘do’ a camera. I suggest that under different circumstances she would talk about ‘doing keyhole’, omitting the head of the noun-phrase *keyhole surgery*. This is a strategy not common in Welsh, and one that doesn’t work in this case, because the phrase she has chosen as the ‘Welsh’ phrase is constructed differently. I suggest that she feels her reference doesn’t work, but fails to find a different reference she considers appropriate.

4. JH: ond hefyd yn yn poeni ynglŷn â a ddylsech chi gael ryw **imiwnedd neu beth bynnag** ag- efo’r fflw moch presennol ma
(‘but also worried about whether you should get some sort of immunity or whatever ag- with this current swine flu’)

JH seems uncertain about how he should express his reference here, as is suggested by his hedging “neu beth bynnag” (‘or whatever’). It’s possible that he’s looking for an expression to refer more specifically to ‘vaccine’ or ‘vaccination’ rather than the more general ‘immunity’.

5. DA: ambell i waith os yd- (.) os oes rywbeth yn bod ar y misglwyf (.)

¹³¹ /ai/ > /ei/ is a common vowel mutation in the morphological process, often accompanying derivational suffixes. Other examples include *haul* (‘sun’) > *heulwen* (‘sunshine’), *naid* (‘a jump’) > *neidio* (‘to jump’), and *mainc* (‘bench’) > *meinciau* (‘benches’).

neu'n amlwg fod yr **hormone** (.) **um merchetaidd** (.) yn annormal (.)
mae (.) chi'n gallu helpu
(‘sometimes if th- (.) if there’s something wrong with the period (.) or it’s
obvious that the (.) um female hormones are (.) abnormal (.) you can
help’)

Merchetaidd is an adjective usually used derogatively of men (‘effeminate’, or more literally, ‘girly’). Dr Ann’s hesitation here suggests that she might consider her reference incongruous, but in the moment fails to think of a different referring expression. Although it is not a particularly common word, the adjective *benywaidd* is more usually used to refer to objects or qualities that are positioned as relating legitimately to women, rather than to qualities that are positioned as non-legitimately woman-like.

6. DA: ond chi'n gallu cael **comesgor**
(‘but you can have a miscarriage’)

Dr Ann uses a verb here (‘to miscarry’) whereas *cael* (‘to have’) requires a noun¹³². Although ‘miscarriage’ is by no means a new concept, there isn’t an easy or commonly used way of referring explicitly and concisely to a miscarriage in Welsh (other than to borrow *miscarriage*)¹³³. I suggest that Dr Ann is familiar with *comesgor* to a degree, but is having difficulty deploying it in spontaneous speech.

Of all disfluency cases, most are Dr Ann’s. There are only two cases by presenters (the ones I present in the examples above) and none I can find by guests. That is not to say that guests do not trip over their words, or pause before making a reference. But the examples here show cases of where the participant has

¹³² *GPC* and *Geiriadur yr Academi* give *esgor* as both noun and verb (as ‘birth’/‘delivery’ and ‘to give birth’, then), but the other codification texts I consulted gave *esgor* and *comesgor* only as verbs. *Geiriadur yr Academi* also gave *comesgor* only as a verb. It is, of course, possible that Dr Ann does view *(cam)esgor* as a noun, although it strikes me as odd. *Comesgor* is not included in *GPC*, whose instalments nearer the beginning of the alphabet were published more than 50 years ago. This suggests that *comesgor* is a fairly recent neologism.

¹³³ Codification texts offer *comesgoriad* (literally ‘miscarriage’), *erthyliad naturiol* (literally ‘natural abortion’) and *colli plentyn* (literally ‘loss of a child’).

committed to a reference. There are more cases in the data of incongruous reference (eight in all) than of mispronunciation (four cases) and difficult reference (three cases).

There is very little repair in the data, either self-repair or other-repair – not in any of the examples above, nor in any of the other cases of disfluency I have identified. Many of the cases of disfluency also do not evidence many other disfluency cues, such as hesitations or truncated sounds (there are three such examples above – examples 3, 4 and 5). To some extent, then, it is my own inference that these are cases of disfluency, and I am aware that I need to take care not to overinterpret here and present speech as disfluent where participants might dispute this. However, I would argue that the presenters and, to a lesser extent, Dr Ann, are trained to carry on regardless, rather than to draw attention to disfluency by retracing their steps. This, I argue, is the main reason why we don't see much repair or hesitations in these cases.

These examples all point to the difficulty of implementing purist ideological discourse, from the perspective of competence, or perhaps of socialisation into institutional norms. Fluency in institutional norms is not given, but acquired, Fairclough (1995a) argues, by the institution's *subjects* (those who have “institutional roles and identities acquired in a defined acquisition period and maintained as long-term attributes” (38), who are “qualified to act through being constrained – ‘subjected’ – to an institutional frame” (39)). Arguably, within the context of the *Syrjeri* data, institutional norms require, to a certain extent, the use of puristic Welsh vocabulary. But in a magazine show like *Wedi3*, that covers every day a variety of different topics, the vocabulary taken from these various fields (each potentially inhabiting an institution of its own) has to be made to fit with S4C's priorities and norms. To a certain extent, this is not particularly problematic in the field of health, since so much of health-related vocabulary is drawn from the lifeworld, as I have argued. On the other hand, more specialised vocabulary is drawn from a world very different from *Wedi3*, one where English

is predominantly the language of specialist communication. The institutional 'subjects' in the *Syrjeri* interaction, cannot be 100% converse in the health-related vocabulary that 'fits' with S4C's institutional norms, since they have not been socialised in its use; the presenters because they are not familiar with the field, and Dr Ann because it is a field she has been socialised into primarily through a different language. Nonetheless, as I have noted above, participants mostly carry on regardless, seeking to give the appearance that talk is seamless, even where some cracks seem to be evident – either where a word is mispronounced, or used 'incorrectly', or where the right expression is not found.

To a certain extent, we can assume that talk varies according to the context of text production. It may seem that the *Syrjeri* broadcasts offer little variation in the way that texts are produced. However, whilst texts come across as fairly uniform, there is some difference in the extent of planning that is carried out for different parts of the broadcasts. There is one example of scripted and edited talk in the data – the pre-recorded clip from *Epilepsy* mentioned above, which is 40 seconds long, and consists of six sentences. That the talk in the *Epilepsy* example is scripted and edited is evidenced, for example, by the lack of hesitations and retracing, the restriction of pauses to comma- or full-stop positions, and the use of dramatic music at relevant moments. Pre-recorded clips are embedded into the live broadcasts. But there are other examples of text that is embedded, having been planned, either fully (like the pre-recorded clips) or partially, beforehand. These include the presenters' introductions that are often read off autocue, especially where the presenter has a number of facts to relate in his/her introduction. But it is clear that the briefing notes written by researchers must be partly embedded into the *Syrjeri* texts, being given to Dr Ann and the presenters beforehand. Often presenters look down on the notes in their laps during the interaction. Dr Ann also seems well equipped to answer the presenters' questions, suggesting that she too has prepared, something she alludes to in *Scarlet Fever*, when she mentions the difficulties she encountered reading up on the condition before filming.

Participants are not tied to directly embedding text from prepared notes, like they are in the case of scripts. But these prepared texts make certain lexical forms available to presenters and Dr Ann from the prior-text (either where the notes have been read before the speech event or where presenters glance at them during the interaction). We can say that these forms are primed – encountered in the prior-text and made available for use, but allowing that in talk they might not be deployed exactly as written.

Prepared text is different from spontaneous speech in that it is usually (as in the case of the briefing notes) prepared through a process of writing which, *inter alia*, allows for more planning and reflexivity. In the case of words that are considered ‘astrus’, in spontaneous speech speakers often aren’t able to access them within the flow of talk, either because they are not familiar with them at all, or because talk moves too fast to allow time for reflection. While all language is reflexive to some extent, in writing we pay much more obviously conscious attention to our linguistic choices (Verschuieren, 2004: 58), in particular in the choice of words. It allows (and often requires) time to be given to consulting codification texts where linguistic forms do not come readily to the author. The *Syrjeri* researcher I spoke to noted that this reflexive process in writing often involved referring to *Geiriadur yr Academi*. In fact, the researcher seemed to find this dictionary particularly useful, rhetorically asking “ble bydden ni heb Bruce?” (‘where would we be without Bruce?’, ‘Bruce’ referring to the main author). Dr Ann also told me that she consulted *Geiriadur yr Academi*. Other methods used in reflecting on lexical form, according to the researcher, were consulting other written texts (specifically a bilingual health-related website) and consulting other members of staff in the open plan office.

How does this kind of writing-specific reflexive process trickle down to the broadcast *Syrjeri* talk? Clear examples of the embedding of scripted text or prepared notes are found in the presenters’ autocue introductions. In these

introductory lines, there are a number of examples of lexical variation of the nature we have already observed, all involving lexical designations for the topics that are under discussion in the day's *Syrjery* item in topicalising clauses¹³⁴:

“cerrig y bustl neu gallstones”; “clefyd y siwgr (.) clefyd melys (.) neu (.) diabetes”; “lupus neu (.) salwch y blaidd”; “sglerosis ymledol neu multiple sclerosis #em #es”; “osteoporosis (.) esgyrn brau”; “y dwymyn goch neu'r scarlet fever”. In another example, two semantically equivalent lexical items are presented in a somewhat more subtle construction which, nonetheless, helps introduce the topic for the day:

“rwan yn ôl y **Meningitis** Trust mae achosion o **llyd yr ymennydd** yn cynyddu wrth i'r clocia droi a'r gaeaf agosau”

(‘now then according to the Meningitis Trust cases of meningitis increase as the clocks turn and the winter draws nearer’)

This clause paves the way for the topicalisation of ‘meningitis’ in a subsequent clause, where the hypernym and cohesive tie, “yr aflwydd” (‘the disease/misfortune’), refer back to *llyd yr ymennydd*¹³⁵. But this construction is more subtle than the others in that the borrowing *meningitis* is used in the organisation’s name, leaving the presenter free to use the Welsh designation by itself (without recourse to reformulation or apposition) to refer specifically to the concept ‘meningitis’. The viewers are presented with the concept ‘meningitis’ in a secondary way, by means of a proper noun containing the lexical item *meningitis*. Since the presenter has already introduced the concept, the Welsh *llyd yr ymennydd* is able to take the position of being designator of the concept ‘meningitis’ other than in a proper noun, whilst its meaning has been clarified¹³⁶.

¹³⁴ “neu” in these examples means ‘or’.

¹³⁵ The clause in full is: “ac yma i drafod yr aflwydd efo ni heddiw mae (guest name)...” (‘and here to discuss the disease/misfortunate with us today is (guest name)...’).

¹³⁶ The use of *llyd yr ymennydd* for ‘meningitis’ isn’t entirely in keeping with the codification texts, which are ambiguous regarding its meaning. Literally meaning ‘inflammation of the brain’, *llyd yr ymennydd* is offered as a semantic equivalent for *meningitis* and for *encephalitis*. Codification texts also offer *llyd y freithell/breithelli* (not found in the data) for *meningitis*. Only in *Terms for*

The significance of these topicalising expressions in the presenters' introductory lines is their centrality, or salience, in the texts. The issues that are under discussion are central to the talk, and since they designate central concepts, participants and researchers are more likely to reflect on the forms used in their lexicalisation. It is arguably of greater importance that the words used to designate these concepts are accessible. At the same time, it could be argued that it is equally important that the words satisfy the purist norms within the institution. In the case of 'lupus', it seems that purism is of such importance that a Welsh designation (*salwch y blaid*, literally 'wolf illness') seems to have been coined for the purpose of the item (as noted above, I have failed to find in any codification text or any place other than in the *Syrjeri* data). Note, however, that this applies only to the designation within the topicalising sequence; in all other lexicalised references to the condition, only *lupus* is used.

Whilst we know presenters' introductions are either scripted (read from autocue) or prepared in some way (adapted from briefing notes or the presenters' own notes), there is some talk in the *Syrjeri* data that either isn't planned or where participants seem to talk outside of the main frame of the interaction. Interaction involving guests tends to open the scope for unplanned talk, largely because guests talk outside of the planned briefing notes. An example of guest-directed talk is found in the following extract from the *Meningitis* text. The presenter, EJ in this case (there are two presenters in this text), has asked the Guest (GU) to describe what happened when she found that her young daughter was ill. We pick up the interaction in the middle of guest's description, where she is relating how she sought to interpret her daughter's symptoms:

Occupational Therapy and in *GPC*, the least prescriptive Welsh dictionary available, have I found *meningitis* listed as a Welsh word (also spelled *meninjeitus* in *GPC*).

Extract 3

1. GU: o'n i'n meddwl taw chest infection achos oedd dim rash o (.) gwbl (.)
I thought that a chest infection (.) because there was no rash at (.) all (.)
2. ar corff ddi o gwbl (.) byddai'n byth uh t- o'n i'n meddwl taw'r
on her body at all (.) I'll never uh t- I thought the
3. **meningitis** (.) yw edrych am rash (.) ond oedd dim rash o gwbl so
meningitis (.) is looking for a rash (.) but there was no rash at all so
4. byddai'n gwith- (.) byth yn gweud //<taw **meningitis**> [?]\
I'll gwith- (.) never say that meningitis
5. EJ: /ond oedd rywbeth\\ yn gweud thoch chi ryw reddf yn gweud thoch
but something was telling you some instinct was telling
6. chi fel mam (.) troeoch chi at eich gŵr a wedoch chi (.) mae hwn yn
you as a mother (.) you turned to your husband and said "this is
7. **feningitis**
meningitis"

We have already seen that *llid yr ymennydd* is the lexical designation that is used to refer to 'meningitis' in the topicalisation sequence at the beginning of this item, although *meningitis* is also present in that sequence. In the intervening interaction, the concept is specifically lexicalised three times, but only with the designation *llid yr ymennydd* (by the presenter, JH, and by Dr Ann). In this extract, we see the guest, however, use *meningitis* (in lines 3 and 4). That is, she does not use the form that seems to have been prioritised by the other participants within the interaction.

Guests don't speak from a script, and they haven't been given a copy of the briefing notes. They are asked to speak from their own personal experiences, and consequently much of the lexical content will be drawn from their own personal histories. *Meningitis*, then, seems to be the form that the guest is most familiar with and favours, despite *llid yr ymennydd* being available from the prior-text and

prioritised by the other participants. Guests are told by production staff that they should feel free to speak in whatever way they want (as I mentioned above). Certainly, this guest's language seems different from the style used by presenters and by Dr Ann. She uses marked borrowings, such as *chest infection* (line 1); and she uses the simplified (widespread but stigmatised) possessive construction *corff ddi* ('her body') without the pronominal pronoun *ei* (line 2). Later in the text, she doesn't use the soft mutation where it might be considered usual ('troi yn du' ('turning black') rather than 'troi yn ddu') and she uses the masculine numeral *tri* ('three') where the feminine *tair* is usual (in "tri oed" ('three years old') rather than 'tair oed'). She seems to show some signs of disfluency, such as the use of a verb form more typically used for the future tense (*byddai'n* (usually meaning 'I will')) where she seems to be looking for a conditional form (such as *bydden i'n* ('I would')) in line 4; she omits a verb from the subordinate clause in line 1 ("o'n i'n meddwl taw chest infection" ('I thought that [it was] a chest infection')); and the complement clause in line 3 doesn't seem to tie cohesively with the subject ("meningitis (.) yw edrych am rash" ('meningitis (.) is looking for a rash')). Her spoken style, then, is markedly different from what is otherwise presented as normal by Dr Ann and the presenters, including, potentially, her use of *meningitis*.

What is most interesting in this extract, however, is that EJ, in her question to the guest in lines 6-7, uses *meningitis* herself. This is despite *llid yr ymennydd* being made available, clarified and seemingly prioritised in the topicalising sequence and in subsequent talk. There are a number of possible interpretations as to why EJ does this. She may assume that *meningitis* is more accessible to the guest, and therefore seeks to make the question easier to understand. Certainly it is important for EJ in this instance that her talk is maximally accessible to the guest, because otherwise the guest may not be able to answer EJ's question, and the resultant interaction may come across as disorderly. It is EJ's job to facilitate orderly interaction. Alternatively, she may have noted the guest's own use of *meningitis* and seeks to converge, perhaps in order to reduce interpersonal

differences between herself and the guest, as might be inferred from the perspective of communication accommodation theory (Giles, Coupland, & Coupland, 1991). On the other hand, looking in more detail at the work EJ is doing here, we see that she uses the lexical item in a quote, “mae hwn yn feningitis” (‘this is meningitis’), where she is assuming the guest’s voice, reenacting the (imagined) moment when it dawns on the guest that her daughter has contracted the disease. EJ’s use of *meningitis* here, then, might be seen to consolidate the shift in voice from her own to the guest’s, on the assumption that this is the word the guest would most likely use in that context – possibly on the basis of the guest herself having used *meningitis*, or considering the guest’s spoken style otherwise. It may be the case, of course, that in so doing, EJ is also reducing interpersonal differences (intentionally or not), by using the guest’s own language and by avoiding a word (*llid yr ymennydd*) that may be assumed to be unfamiliar to the guest. The use of linguistic resources that are not mutually intelligible necessarily creates social distance. Whatever the most fitting interpretation (if there is only one), purist ideology and planned language (if *llid yr ymennydd* is planned language – see note 136) are not dominant in this case¹³⁷.

A similar example is found in *Lupus*, where EJ repeats the guest’s words, including the noun *kidney*. The guest is in the process of relating her experiences when she became ill with lupus (*SLE* is a form of lupus):

Extract 4

1. GU: wen i di cael yr #es #el #ee wedyn o’n i di cael y systemic oedd e di
I’d gotten the SLE then I’d gotten the systemic it had
2. effeithio pob organ (.) ond y kidney
affected every organ (.) except for the kidney
3. EJ: ond y **kidneys** =
except for the kidneys

¹³⁷ Notice, however, that EJ does mutate *meningitis* (correctly, according to the norms of ‘standard’ Welsh). Borrowings, especially those that are less common or that are marked, are not always mutated in the same way as ‘native’ Welsh words. But EJ’s mutation here does help naturalise the use of a borrowing here within the dominant purist frame.

4. GU: = ond y kidneys
except for the kidneys

EJ is again, in a sense, voicing the guest's words here, although not in reported speech like in the *meningitis* example. But again she is repeating the guest's lexical use, using a word, *kidney*, that is different from Dr Ann's use in the preceding text (Dr Ann uses *aren* twice), and that seems to run contrary to purist norms and to the stipulations of lexical planners¹³⁸. In this case, however, since EJ is immediately echoing the guest, to use a different lexical form, namely *aren*, would likely be perceived as a correction, a face-threatening act that EJ is actively seeking to avoid in the pursuit of orderly interaction. She is not going to coax guests into relating their stories by correcting their language. We see here, then, that orderly interaction requires laying aside purist ideology in favour of accommodating guests. (Interestingly, later in the text, both Dr Ann and EJ use *aren*, in one instance (by EJ) in directing a question to the guest.)

Such examples of guest-oriented talk, however, are fairly rare in the data, most of the talk by Dr Ann and the presenters revolving around the topics that are covered in the briefing notes – introduction, symptoms, treatment, and advice, as the *Syrjeri* researcher related to me. Even in this talk, however, there are cases where the lexical items used do not seem to follow purist norms. In a number of such cases, however, discursive devices are used to move the boundaries of the interactional – and institutional – frame. There are a number of cases where lexical designations are explicitly ascribed to other people or other contexts, by means of constructions such as “beth ni'n galw X” (‘what we call X’) or “X fel ma nhw'n gweud” (‘X as they say’). I have counted approximately 40 cases of such metalinguistic marking in the *Syrjeri* data (an average of two in every item, then). In most of these cases, the quotative verb (and I argue we can talk about quotatives, since they, in a sense, introduce somebody else's words) is in the

¹³⁸ *Kidney* is more obviously disliked by codification texts than *meningitis*. Not even the more inclusive *GPC* accepts *kidney* (or *cidni*) as Welsh (except for in the phrase *cidnibêns*).

present tense. These constructions, then, overwhelmingly refer to ongoing, habitual language practices. Two examples, from the *Bunions and Corns* item are presented below. In the first, Dr Ann is describing what causes a bunion:

ac oherwydd hynny (.) mae na bwrs: (.) ychwanegol (.) annormal (.)
yn ffurfio wrth ochr (.) tu fewn (.) bys bawd (.) y droed **beth mae**
nhw galw bursa
(‘and because of that (.) an additional (.) abnormal (.) purse (.) forms
near inside the big toe what they call a bursa’)

In the second, she is suggesting a way a bunion can be dealt with (showcasing the shoe she is referring to):

mae rai bobl yn ffindo (.) y sgidie hyll ofnadwy ma os gai weud (.) yn
help **beth mae nhw galw (.) #em #bi #ti_s**
(‘some people find (.) these incredibly ugly shoes if I may say so (.) a
help what they call (.) MBTs’)

In both these examples (and in most of the other similar cases of metalinguistic construction), Dr Ann is focusing attention on the lexical designations with heavy stress on the first syllable and a pause – a marked pause in the second example and delayed release of /b/ in *bursa* in the first. Focusing through emphasis and/or pausing is a strategy that is used extensively by Dr Ann in her scientific expositions, as we see in the other words in the first example: *pwrs*, *ychwanegol*, *annormal*, *tu fewn*, *bys bawd* are all heavily accented. But the focusing on *bursa* and *MBTs* is also accompanied by these metalinguistic constructions, “*beth mae nhw galw...*” (‘what they call...’).

Notice that the verb subjects in these constructs are personal pronouns, in these cases 3rd person plural *nhw* (‘they’). This is, in fact, the most prevalent verb subject in constructs of this type (occurring 17 times). These quotatives, then,

attribute the language practices thus framed to unspecified third parties. What Dr Ann seems to be doing here is to place the word under focus outside of the interpretive frame of the interaction. In his definition of genre, above, Gumperz (p.136) talks about *bracketing* talk – a way of placing talk within a particular frame. But there are frames other than generic – sociocultural and interactional (Coupland, 2007) as well as, I would argue, institutional. Dr Ann is still focusing on the lexical designation through emphasis, like she does with the other lexical content words in the first example. But by framing the lexical designations (*bursa* and *MBTs*) in this way, Dr Ann is taking these words out of the bracketed talk of the current interaction. Thereby, these words do not have to be interpreted in the same way as the rest of the talk, and consequently, Dr Ann is able to use them. This bracketing, transient as it is (covering no more than a single lexical item), licenses the use of words that would normally be outside normal institutional practice. In all such cases, however, semantic meaning is clarified (in the examples above, by means of a visual aid and a prop).

Why does Dr Ann place these words outside of the interpretive frame of the interaction¹³⁹? It may be because they are borrowings. Placing a new boundary around these words, reframing their contexts of use, can allow borrowings to be used where they may not be accepted otherwise. This practice can be seen as a spoken equivalent of italicising or scare-quoting borrowings in writing. Alternatively, it may be because they are specialist words – Dr Ann is drawing attention to the fact that viewers won't be familiar with these words. These words are not presented as part of the *Wedi3* voice, the voice that is supposed to be on an equal footing to that of the viewers engaging in this informalised friendly chat within the realms of the lifeworld. These words are presented as belonging to the medical world. The 3rd parties for whom these words are a part of ongoing habitual practice may be unspecified, but invoking these unspecified 3rd parties within the context of medical talk, and, in the case of *bursa*, whilst

¹³⁹ I talk about Dr Ann here because most of the metalinguistic constructs of this kind are hers. However, there are a small number by other participants, including one by a guest.

showing a full-screen x-ray of the human foot, suggests that Dr Ann is alluding to medical experts. In fact, in a number of cases (eight in all), Dr Ann includes herself within the scope of reference, using the 1st person plural personal pronoun *ni* ('we'), such as in Extract 2 above (line 3), "beth ni'n galw anaphylaxis" ('what we call anaphylaxis'). Here Dr Ann is invoking her identity as a professional medical expert and voicing *anaphylaxis* as her own word, albeit not a word that can be freely used within the context of the *Syrjeri* interaction.

Whatever Dr Ann's intent here, her framing of these words can be interpreted as licensing either the use of borrowings or of specialist words. Similar to the use of italics and single quotation marks in writing, it can serve more than one purpose. Examples in the data of non-specialist and non-borrowed words framed in such constructions are very rare. I have only found *therapi deall ac ymddygiad* ('cognitive behavioural therapy'), although it is a fairly literal translation of the English; and *cytshin* ('catching'/'contagious', pronounced by EJ like English *kitchen*), in fact a borrowing, but thoroughly integrated and framed by EJ as an old dialect word¹⁴⁰. Almost all of the words that are highlighted in these types of metalinguistic, reframing construction, then, are specialist and can be interpreted as borrowings (even those of classical origin can be interpreted as borrowed via English): *adhesions, anaphylaxis, angina, atypical, booster, bursa, butterfly rash, bypass, cholecystitis, chronic bronchitis, coronaries, electroencephalograph, fibrillation, fit, grand mal, keyhole surgery, MBTs* ¹⁴¹, *MRI scan, orthoses, patch tests, Pel Ebstein virus, plaques, podiatrist, strawberry tongue, synapse, systemic lupus erythematosus*. In fact, in the case of the latter (from the *Lupus* text), Dr Ann explicitly highlights the status of the lexical item as a borrowing from (or via) English whilst explaining in some detail the meaning of this specialist phrase:

¹⁴⁰ In fact, it is the only example in the data of a reframing construction that refers to lapsed practice, using the imperfect tense: "pa mor (.) cytshin o'n ni'n arfer gweud" ('how (.) cytshin we used to say').

¹⁴¹ In the case of acronyms, Dr Ann uses English labels for the letters of the alphabet (e.g. /ai/ not /i/ for the letter 'i'), as well as, in the case of *MBTs*, the English plural suffix /z/.

felly ni'n galw fe (.) yn Saesneg (.) **systemic** (.) achos bod e'n (.) bob
 system (.) **lupus** (.) achos (.) ffurf (.) y brech (.) **erythematosus**
 erythema yw jyst gair arall am coch (.) brech coch
 ('so we call it (.) in English (.) systemic (.) because it's in (.) every
 system (.) lupus (.) because of (.) the form (.) of the rash (.)
 erythematosus erythema is just another word for red (.) red rash')

What we see here is the voicing of a medical world that is lived primarily through the medium of English. Dr Ann uses these words because they are the words that have currency amongst medical practitioners. She is talking to an audience who might come into contact with the medical world and uses the labels they might encounter. The *Syrjeri* items are meant to be partly informative, to disseminate knowledge that might be of use to viewers in the real world, were they to find themselves face-to-face with the medical world in their own personal experiences. In fulfilling this institutional priority, Dr Ann needs to use language that will serve the viewers well in a real-life situation. Indeed, some guest talk evidences that 'ordinary people' do come across these labels, and in picking them up they become meaningful to them. They appropriate rather than translate them, although they are slotted into Welsh syntactic structures. Guests typically use the specialist designations they encounter, such as: *keyhole surgery*, *puncturo'r bile duct* ('punctured the bile duct'), *y long-acting* ('the long-acting (insulin)'), *peripheral neuropathy*, *Hydrocortisone Cream*, *eli one per cent* ('one per cent cream'), *intensive care*, *rehab*, *suppresso* ('to suppress'), *relapsing-remitting MS*.

We are reminded here that language does not merely reference objects in the world, but also references itself – not only through overtly metalinguistic commentary, but also through the *repetition* of linguistic forms over time. Dr Ann's repetition of the habitual (oft repeated) talk of medical practitioners is an intertextual link between her current practice and their previous practices (which she frames as ongoing). The same is true of the above examples of the

guests' talk, although less overtly marked. The difficulty for corpus planners is that the intertextual link between current practice and the previous practices it refers to is obscured through translation. If repetition creates and draws on a history for the element repeated (Johnstone et al., 1994: 12), how direct is the intertextual link between the repetition in translation and the original element repeated? It is not that there is no intertextual link at all. In fact, I argue that much of Welsh lexical planning relies on Welsh speakers being able to retrieve the original form through translation in order to make sense of 'Welsh' 'coinings'¹⁴² (e.g. to retrieve *cognitive behavioural therapy* from, and to make sense of, *therapi deall ac ymddygiad*). Brown (2000) emphasises that repetition varies in degree; that often both form and meaning are repeated, but that equally, repetition can entail repeating only meaning or only form. But arguably, where both form and meaning are repeated, the repeated element is more easily accessed. Using planned lexis can, then, require the speaker to use language where intertextual reference is obscured¹⁴³. Arguably, then, until planned language achieves some currency, outside of contexts where purism and standardisation are dominant ideologies speakers may not prioritise the use of planned talk over transparent intertextual reference. Paradoxically, however, it is only through use that planned language can achieve currency, currency being another way of talking about repeated ongoing practice.

Conclusions

The recurring theme, I feel, in this discussion is that purism is a prevalent ideology, or ideological-discursive formation, within the interaction, but only to the extent that it does not clash with other institutional priorities, such as the need to inform and entertain, by means of Dr Ann's and guests' narrative, or the need to link directly and transparently to the medical world referenced. In cases where purism does potentially clash with other priorities, sometimes both purism and the clashing priority are accommodated side-by-side, as in the case of

¹⁴² My scare quoting of 'coining' here is to question whether a translation ought to be considered a coining at all.

¹⁴³ Not to mention the mental gymnastics necessitated by in situ translation.

the reformulating appositions (but often only in cases where reflection on appropriate lexical forms has been enabled by the process of preparing for the broadcast). In other cases, such as the metalinguistic constructions that shift the interpretive frame, the suspension of purism is licensed by taking what is said out of the bracketing of the current interaction. However, there are cases where borrowings are used in normally-presented talk, such as *jaundice* in Extract 1. Another example found in the data extracts presented above is *rash* (Extract 2). Notably, I have previously categorised both of these terms as not codified. What of the terms from my codification tables? In fact, only nine of the not codified terms, out of a total of 26, are presented as normal in the interaction (at least, in the texts used for the sample). Of these, a smaller number are borrowings (*autoimmune*, *glandular fever*, *jaundice*, *keyhole*, *straen* ('strain (of virus)') and *sic*). The others are either presented as not normal (used as a gloss or attributed to others) or they are used by guests or towards guests. It may be that guest talk is also presented as normal. However, arguably guest talk does not represent the voice of S4C. There are only very few terms in the data, then, that represent S4C's voice, and that clearly contravene the stipulations of codification texts.

Note, however, that in the qualitative analysis I have primarily emphasised *purism* rather than standardisation or planned language. I have done so because it seems that purism is the prevalent ideology at work in the institution, and not standardisation (the suppression of variation) per se. This is an observation that I have made in both parts of the analysis, in fact. Although staff at the production company did refer to the informal policies they followed when considering appropriate subject lexis in preparation for the interaction (e.g. consulting *Bruce*), language practices do suggest that finding overtly *Welsh* forms (e.g. *salwch y blaidd*, *bustl*) is given greater priority than finding forms authorised by lexical planners (*lwpws*, *coeden y bustl* ('gallbladder')). I suggest that purism is a prevalent ideology in the *Syrjeri* interaction, but that it clashes with other institutional priorities. Formulated in the terms of the LPP literature, purism is a

covert language policy, but its implementation is challenged within this particular context by the other demands of the interaction.

CHAPTER SIX:

Work talk at the Housing Association

6.1 INTRODUCTION: THE HOUSING ASSOCIATION DATA

As in the previous analysis chapter, my aim in this section is to present an overview of the data, in order to familiarise the reader with the data, to outline the content of talk as well as to introduce some detail in terms of participant relations necessary for the subsequent analyses. I start by outlining the structure and content of the meetings, before discussing what I infer to be discourse goals and priorities, in particular in relation to the institution's broader aims. There is a sample transcript (a short section from the *Voids* recording) in Appendix 3.

Structure and content of Housing Association meetings

The four meetings recorded at the Housing Association vary in their length, from the shortest at 53 minutes (*Estates*) to the longest at 2 $\frac{3}{4}$ hours (*Minafon*). Meetings also vary in the number and type of participants involved. They all involve Housing Association staff, but these vary in their positions in the organisational hierarchy and hence their responsibility, or executive power, within the organisation (see Chapter 4 on the hierarchical structure of the Association). In *Estates*, which is a meeting of the *Tenancies Team*, there are 17 participants, all of whom are staff of lower-level responsibilities, other than the chair of the meeting, Phoebe, who has a mediating role between the senior management team and staff of lower-level responsibility. In *Voids*, a meeting of the *Voids Working Group*, there are nine participants, including a highly ranked member of staff, Nell (the chair of the meeting), who encourages debate but strategically leads the discussion. In *Minafon* there are eight participants, many of whom are staff of higher-level responsibility, including two highly-ranked members of staff, namely Phil, who mostly directs the discussion, and Nell. There is also one external consultant in this recording. Finally, *Publications* is a meeting of the *Publications Panel*, which consists of four Housing Association staff and five

tenant representatives (nine participants in all). The *Publications Panel* is one of a number of tenant panels that the Association holds as a way of giving tenants a voice within the Association.

All meetings took place in the Association's committee room at their main premises, and all had a very similar set-up. The room was laid out and furnished specifically for holding meetings – a collection of tables and chairs in the middle of the room put together to sit approximately 16 people, some notice boards on the walls, a white board and projector. In some meetings a laptop computer was hooked up to the projector, and in another meeting a flipchart was placed close to the table. In short, it was a fairly typical and functional committee room. Participants sat around the table, where they would usually stay until the end of the meeting. One of the participants acted as chair in every meeting and another as minute-taker. These were the only formally designated meeting roles that are consistent across all recordings. In some meetings, individual participants were asked to lead or give a presentation on specific agenda items. There was a simultaneous interpreter in the meeting of the *Publications Panel* in order to allow for the meeting to be conducted bilingually¹⁴⁴. The interpreting facilities were used by three of the tenants.

Typically, the chair of each meeting leads the discussion, welcoming participants (and myself), referring them to the relevant documentation, and leading from one item of discussion to the next. In most meetings, there is a formal written agenda, which the chair uses to structure the meeting. The meetings, then, can be seen as speech events, made up of shorter events, or episodes, each focused on a specific agenda item, and related to the broader topic of the meeting. Only in *Estates* was there no formal agenda, and, coupled with the participants' lack of executive power, the effect is a discussion that tends to go in circles with no clear

¹⁴⁴ The interpreter sat aside from the participants, whispering the interpretation into a microphone, which transmitted the message to headsets worn by those who required interpretation. Only Welsh-language utterances were interpreted.

goal or outcome¹⁴⁵. Even in the case of the other three meetings, structure isn't absolutely rigid. Sometimes agenda items are swapped impromptu, participants other than the chair ask questions, talk sometimes diverges into two or more different discussions, etc. But the chair always reconvenes participants to a single discussion. There are also some 'unofficial' episodes, or asides, where talk veers off the issue currently under discussion or the broader topic of the meeting, or very occasionally even off the Association's business altogether.

Talk in the meetings covers a number of different topics, related to the broad goals of each meeting. These broad goals can be characterised as: going through lists of properties to discuss which are vacant, which need action, etc. (*Voids*); various steps that need to be taken to move forward a housing development (*Minafon*); discussing and deciding on the best format for a tenants' handbook and other materials and ensuring tenant input on this (*Publications*); and deciding on the best way to proceed after a tenants' consultation (*Estates*).

It is difficult to generalise about the topics covered in the pursuit of these goals, due to their varied nature, but some examples are: organisational systems and procedures for letting properties (*Voids*); the condition and attributes of properties (*Voids* and *Estates*); the design of new property and surroundings (*Minafon*); project management and scheduling (*Minafon*); rents and charges (*Minafon*); installing solar panels (*Minafon*); lay-out and look of publications (*Publications*); liability and insurance in the case of damage to property (*Publications*); how tenant complaints and queries are logged and dealt with (*Estates*); getting new kitchens fitted (*Estates*); budgets and procedures for maintenance and upgrading work (*Estates*).

Discourse goals and priorities

Talk mostly orients to pursuing the agenda items, seeking to decide on the most

¹⁴⁵ That is not to say the meeting had no use – it did seem to have the effect of consolidating the group and for participants to share knowledge, if informally.

appropriate future action. Typically, discussion of agenda items is strategic – with participants discussing issues, giving background information, gathering and sharing information, highlighting or identifying problems, seeking solutions, determining future action and allocating responsibility for action (although these goals may not be salient to participants, and participants vary in the degree to which they orient to strategy). Such talk derives from and contributes to achieving the overall aims of the Association. These aims are identified in its corporate plan, a formal policy document written and approved at the highest levels of the organisational hierarchy. The Association's aims are broken up into various strategic objectives, which are then broken down further to tasks and targets for whose implementation individual staff members are responsible. Trickling through the organisation there is a chain of command, with each member of staff responsible for his or her own set of objectives and tasks, which together is supposed to work towards the implementation of the Association's overarching aims – and through the aims, its 'vision'. Meetings facilitate the pursuit of these aims and strategic objectives, through collaboration and sharing information. But they are also one means of continuously monitoring whether aims and objectives are being met, and of providing a written record (through the minutes of the meetings) to attest that this monitoring and checking is taking place. By some measure, meetings are performance-checking devices – checking the performance of teams and of individuals. Talk in meetings that is oriented to the strategic pursuit of the Association's aims, then, can be called *institutional* talk. In institutional talk, participants play their institutional roles, framed by the requirements of the institution. Institutional talk in meetings results in action being taken outside of the meeting, usually noted in the minutes, the official record of what was said (i.e. within the institutional talk) and what was decided.

Within institutional talk, then, there is the overall aim of moving the work of the Association along, through the work of individual members of staff in collaboration with each other. Typically, it is the role of the chair to move the discussion along in the right direction, although in some cases (such as in

Publications) this role is shared. In *Voids*, for example, we see that the chair, Nell, is quite strict in how much she allows talk to veer off-topic. For example, when talk within the group splits up into two or more individual discussions, she brings everybody back together, noting that the 'recorder is recording about four different conversations' (possibly using the recorder as an excuse). Elsewhere, when somebody leaves the room to answer their mobile phone, she says to the rest that she hopes the call is important, warning them not to do the same.

Although most of the talk can be described as institutional, it should not be forgotten that there is other talk in the meetings, mostly at the peripheries (meaning that such talk is not the main focus of attention, although some of this talk also appears at the beginning and end of the meetings). Other talk might be completely unrelated to the institutional goals pursued in the meetings, or might be a tangent from those goals. Participants, then, might play their institutional roles to a greater or lesser extent (although, since the meetings all take place on the Association's property, staff members always inhabit their institutional roles to a certain extent). I also emphasise that participants seem to differ in terms of their concern with pursuing institutional goals. Whilst all participants are responsible for carrying out their own formally specified tasks within the Association, their differentiated ranking within the Association impacts on their responsibility for and interest in the wider priorities of the Association. Staff members that are more highly ranked (as managers) have more of an interest in ensuring that goals are met, and therefore that talk in meetings remains on-topic. In their institutional roles, higher-ranking staff members have a greater responsibility for safeguarding the institution's interests, rather than merely carrying out formally-designated tasks.

6.2 SUBJECT VOCABULARY AND CODIFICATION TEXTS

Considering the wide scope of topics covered in the meetings, it is difficult to generalise about the vocabulary content of the data. However, there are some categories that seem to be salient. We find vocabulary that is specific to the

housing-related work of the Association – that is, vocabulary related to houses and property, their content and how they are processed by the Housing Association. These include, for example, vocabulary that categorises different types of property (e.g. *bungalow*, *fflatia dau berson* ('two-person flats')); vocabulary items to talk about letting property (e.g. *eiddo gwag* ('vacant property'), *notice*, *statement rhent* ('rent statement'), *ailosod* ('to relet')); and vocabulary related to property features, fixtures, construction and maintenance (e.g. *soundproofio* ('to soundproof'), *underlay*, *cooker*).

But we also find vocabulary from other fields that are related to business, or business-like organisations (the Housing Association is a business, but one that is driven by a social objective, rather than profit-making). These include vocabulary related to finance and financial management (e.g. *grant*, *gwariant* ('expenditure'), *gwir gosta* ('true costs'), *equity*, *arbedion* ('savings'), *buddsoddiad* ('investment'), *rheoli cyllideb* ('budget management')); vocabulary related to laws and regulations (e.g. *health and safety*); organisational management and procedures (e.g. *feasibility*, *rhestr fer* ('shortlist'), *contingency*); and general business dealings, (e.g. *contractor*, *logistics*, *lansio* ('to launch'), *supplyio* ('to supply'), *specification*, *ymgyngori* ('to consult')).

Note that most of these vocabulary items, including the items specific to housing and property, are borrowed (sometimes translated) from practices and working cultures external to the Association (although, the Association is embedded within these cultures), such as from the development of knowledge and thinking about finance and financial practices. Business organisations typically employ staff with expertise from these other worlds, such as those who have some formal training and experience (socialisation) in finance. They bring these worlds with them and build them into the organisation of the business. Vocabulary specific to housing is also borrowed in some way from outside the organisation. Clearly *underlay* comes from the world of carpeting, and *cooker* from the everyday experience of anybody who uses a kitchen.

Moreover, the conceptual constructs that these vocabulary items invoke, although largely related to the work of the Association, are not particularly specialist, since they are largely used and understood in the wider society in which the Association is embedded. For example, although *eiddo* ('property') frames land and buildings in a particular way, as being in the legally warranted and enforced ownership of particular persons or entities, this way of viewing land and property is fundamental in the prevalent culture in which the data is embedded – not only the culture of the community of practice, but the culture within the political borders of the United Kingdom, and even dominant global culture. Every child within this cultural space is socialised into having a seemingly innate, if seldom consciously rationalised, understanding of property. Although *eiddo* is a central 'term' in the Association, it is not specific to it, and *eiddo*, when used in interaction within the Association isn't used with a semantic meaning very different from that used otherwise.

But there are some vocabulary items that *are* quite specific to the Association. As a community of practice (see Chapter 2), we can postulate that the staff of the Association have developed, over time, a shared repertoire of resources, including vocabulary, for making meaning in engaging in their shared endeavour (Holmes & Meyerhoff, 1999). These include the items borrowed from other worlds and wider society which have become central to the Association's organisation, but there are some vocabulary items that refer to Association-specific constructs, its specific objects of knowledge (Goodwin, 1994). Mostly these objects of knowledge relate to its internal organisational systems and structures. For example, the Association's procedure for dealing with tenant notifications of faults in properties (as explained in *Estates*) consists of labelling (or 'coding' in Goodwin's terms) different phenomena as objects within the system. When a tenant notifies the Association of a fault (in writing or verbally) at a property that needs attending to, this notification is called a *cwyn* ('complaint'). The communication, now coded as an object – a complaint – within

the system, is officialised, and the Association commits to respond to it (when the system works optimally). The member of staff who received the notification *riportio* ('reports') it, which means creating a record of the complaint in the central computer system, and turning the *cwyn* into an *archeb* ('order'). Then the responsible member of staff decides the priority level of the order – either *brys* ('urgent'), *saith diwrnod* ('seven days'), *tri-deg diwrnod* ('thirty days') or *ar suspend* ('on suspend'). And on the system goes. Thus, whilst all of these vocabulary items are borrowed from other worlds – in fact, from quite general/basic vocabulary – they have very specific meanings within the Association's system of dealing with complaints. Along the line, phenomena are conceived of as *cwyn* or *archeb* or *archeb ar suspend*, all of which are to be understood in the light of the Association's system of complaints. Moreover, these phenomena typically have fixed labels within the Association.

Other vocabulary items that refer to Association-specific objects of knowledge are division names, such as *Cynnal a Chadw* ('maintenance') and *Tim Tenantiaethau* ('tenancies team'); job titles or scheme names (the latter two I avoid naming explicitly for reasons of anonymity); as well as names for specifically allocated budgets (*cyllideb ail-leoli* ('relocation budget')), policies (*strategaeth rheoli asedau* ('assets management strategy')), legal documents (*cytundeb tenantiaeth* ('tenancy agreement')), etc. Such objects are central to the organisation of the business, and their labels typically have an official flavour.

Nonetheless, sometimes some of these objects are referred to with shorthand labels. For example, *Cynnal a Chadw* above is a shortened version, using only the verb-phrase to refer to the team/division. Similarly, *cynlluniedig* ('planned'), an adjective, is used to refer to a budget allocated for carrying out planned upgrading and maintenance work on properties, as well as to the list recording such work. Other objects are referred to with different labels. For example, in *Voids Nell* uses the reference *criw Datblygu* ('the Development crew'¹⁴⁶) to refer

¹⁴⁶ *Criw* doesn't have the 'gang' or 'posse' connotation that English *crew* does.

to the development division, emphasising the people who work there rather than presenting the division as a non-human entity. And in *Estates*, participants refer to a specifically allocated budget as a *pot*, conjuring up the image of a tangible object where tangible money is physically kept, presenting budgets as having a physical reality, as in the following example by Lauren:

“raid i (personal name) roid pres yn y pot i ymweliada stad”
 (‘(personal name) has to put money in the pot for estates visits’)

Participants in this recording also refer jokingly to a seemingly bottomless *pot i neud sandwiches* (‘pot for making sandwiches’), that is, money allocated for the provision of sandwiches and other refreshments in meetings.

It should be noted that not all of the types of vocabulary items discussed so far can be assumed to be known to all staff or dealt with by all staff members on a daily basis. Staff members who are responsible for finding tenants for properties, for example, are unlikely to have to talk very much about cookers and carpet underlay, which is largely the territory of those who work for *cynnal a chadw* (‘maintenance’ (division)) and the *gweithlu* (‘workforce’). They, in turn, do not have to talk much about rent statements and notices. However, all are likely to be familiar with labels referring to the objects most central to the organisation of practices at the Association, such as its divisions and higher-level staff.

Finally, despite being able to identify broad fields from which a number of lexical items are drawn, there remain a significant number that are not easily categorised, yet they are central to the way participants communicate in the data. Vocabulary items such as *issue* and *sortio* (‘to sort (out)’) can be difficult to group according to work-related or Association-specific categories, and yet it is hard to dismiss such vocabulary as obviously peripheral to communication in the workplace. Both *issue* and *sortio* occur relatively frequently in the sample (see Appendix 5), more often, in fact, than some housing-related vocabulary. Compare

issue (used with the meaning ‘matter’ or ‘problem’) with *threshold*, a housing-related word. *Threshold*, a word we find only in *Minafon*, occurs only twice, the second being a clarification of the first, where Gaby has had difficulty remembering the word. Gaby uses *issue* a total of ten times, on the other hand, and it occurs 47 times in all across the transcribed data, coming up in all four recordings. Surely, *issue* must be considered more central to workplace communication. But it is difficult to know how to identify the specificity/centrality of words of this nature to workplace interaction, other than by frequency counts.

In terms of selecting vocabulary for inclusion in my quantitative analysis of lexis in the data as codified, I have sought to include lexical items from the fields that I infer to be related to the work of the Housing Association, such as housing and business. I have taken a wide scope of fields as being related to the business of the Association, since they deal with a number of fields in the pursuit of their business – including construction, social services, local authorities, the courts, etc. I have also sought to include the Association-specific objects of knowledge (or their labels) that I could identify. As I noted in Chapter 3, however, lexical items of this nature – Association-specific – are less likely to be found in codification texts precisely because of their specificity. Some of these objects are salient in working practices outside of the Association, and might be prevalent enough for lexical planners to warrant including in a codification text (e.g. *cytundeb tenantiaeth* (‘tenancy agreement’)). But others, including shorthand or alternative versions, are less likely to be codified, despite not necessarily going against the working principles of lexical planners. They do, however, tend to be built from resources that are considered by planners (typically being multi-word items). In terms of the lexical items, such as *issue*, that I have found difficult to categorise, I have sought to include those that I felt to occur frequently. As with the *Syrjery* data, participants do make use of metaphor. I have not typically included these, unless they seem to occur quite frequently and used as quite salient objects of knowledge.

After deciding on the terms to be included in analysis, I considered which codification texts would best be consulted to investigate the degree to which these terms can be said to be codified, as I did with the media dataset. Again, I consulted the general codification texts outlined in Chapter 3, but alongside these I consulted some texts more specific to the vocabulary that emerges in the data. Housing isn't a field that has been the focus of any specific lexical planning effort – there are no glossaries of housing terms. Moreover, the topics covered in the workplace data were less focused on a single field than in the media data. Nonetheless, the various fields that emerge in the data are, to some extent, catered for in lexical codification texts, especially those fields specific to most businesses or business-like organisations. These include, for example, finance, project management, organisational structure, legal matters, etc. Moreover, the generally less specialised, and more culturally embedded nature of the talk, as I discussed above, suggests that the general codification texts may be equally relevant.

I have focused on the relevant texts available in the *Cysgeir* application, because of their usability and currency. These texts, although fairly marginal to the Association's work, include a number of more generalised lexical items. These texts include¹⁴⁷:

- *Environment Agency Wales Terms* (Prys, 2002a)
- *Dictionary of Finance Terms for the National Assembly of Wales* (Prys, 2000a)
- *Terms in the Health Care of Older People* (Prys, 2005)
- *Social Work and Social Care Terms* (Prys, 2000b)
- *Terms for Occupational Therapy* (Prys & Davies, 2007)

Other texts might also have been relevant, including *Termau Adeiladu*

¹⁴⁷ I give only the English titles here. Welsh titles can be seen in the list of references.

(‘Building/Construction Terms’) (Griffiths, 1993) and *Termau Llywodraeth Leol / A Glossary of Local Government Terms* (D. G. Lewis, 1996). However, I didn’t consider that there was much to gain from including them, given that they were not central to the work of the organisation, fairly short and not digitised, hence less user-friendly. The amount of work that each further codification text added to the workload meant that it was better to be selective in the texts I included. Considering that, as we shall see, there are very few terms that are coded ‘not applicable’ (i.e. there are very few terms that were not found in any of the codification texts), the codification texts that I did consult were typically relevant enough.

6.3 CODIFIED LEXIS AT THE HOUSING ASSOCIATION

The sample I selected for this part of the analysis consists of sections, of similar length, selected from each of the four recordings. These make up over 30 minutes of talk. In this sample I categorised 180 terms as relating to the work of the Housing Association, and 472 individual instances of terms. Table 6.1 shows the distribution of term types and tokens according to the degree to which they can be said to be codified (as discussed in Chapter 3)¹⁴⁸. As in the case of the media analysis, I categorised each term as codified or not, according to how it was dealt with (if it was dealt with at all) by the codification texts. In a number of cases, however, it wasn’t possible to categorise a term simply as codified or not, and the final categories I used are as follows: codified, not codified, partly codified, mixed, and not applicable. There is a detailed explanation of what these categories mean in Chapter 3, although I briefly explain below, where I discuss each category separately.

Considering first the term types, we see that the category with by far the highest percentage of term types is the codified category, with 108 types, or making up 60% of all term types. Well over half of all term types, then, are codified unambiguously by the texts consulted as Welsh words. There are a number of

¹⁴⁸ Appendix 5 shows all the terms found in the sample, listed according to codification category.

borrowings (26 in all) in this category, but none that are saliently non-native; all borrowings in this category are either integrated or have no linguistic features that could be integrated. The former type includes, for example, *asesiad* ('assessment'), *clerc* ('clerk'), *lansio* ('to launch'), and *plismon* ('policeman'), as well as some potential loan translations, such as *rhestr aros* ('waiting list') (although it is debatable whether such lexical items ought to be considered borrowings). Borrowings in this category that have no linguistic features that could be integrated (other than possibly some marginal phonetic integration) include *panel* (in both senses found – see Appendix 5), *drafft*, *cost*, *bin* and *system*¹⁴⁹.

Category	Types	%	Tokens	%
Codified	108	60.00	331	70.13
Not codified	45	25.00	81	17.16
Partly codified	5	2.78	11	2.33
Mixed	18	10.00	43	9.11
n/a	4	2.22	6	1.27
Total terms	180	100	472	100

Table 6.1: Types and tokens of all terms in the Housing Association sample

A small number of the items in this category might be described as basic vocabulary – words we expect participants to be familiar with from quite young. These include, for example *tŷ* ('house'), *cegin* ('kitchen') and *llawr* ('floor'). They have been included because they are related to properties, however, and it is no great surprise that there are fairly few other examples of basic vocabulary in the data. Nonetheless, as I explained above, much of the vocabulary found in the data is fairly general – the concepts they denote do not require specialist knowledge to understand, largely because they are quite prevalent in the more widespread

¹⁴⁹ *GPC* lists *panel* as an early borrowing from Old French, but in the sense of a piece of cloth placed under a horse's saddle. Both senses found in the data (a sheet of material and a group of people) are semantic extensions of the original word. This process of semantic extension might arguably have been borrowed from English. *GPC* lists *cost* as a borrowing either from Middle English or from Old French.

culture. Many of these items do refer to concepts that are quite central to the organisation of the Association. Examples found in the codification category are *archeb* ('order'), *enwebiad* ('nomination'), *gosod* ('to let') and *llawlyfr* ('handbook'). We expect, however, that participants are familiar with the concepts (if not necessarily their codified Welsh designations) from outside the organisation. Nonetheless, being items that are central to the organisation of the Association, it is not surprising that the designations are codified as Welsh – being semi-institutionalised in the context of the interaction, they receive fairly formal designations, and these designations tend to be repeated in organisation-internal interaction. It is notable, however, that a number of the items in this category that can be described as Welsh-origin have borrowed variants elsewhere in the data, including *archebu* (*ordro* > 'to order'), *cyflenwr* (*supplier*) and *llawlyfr* (*handbook*). It is possible that there is some contextual sensitivity here – that the more codified, potentially institutional designations are used where participants are working towards the priorities of the institution, and that their borrowed variants (typically not codified) possibly index a discord with the institution.

Table 6.1 also presents the number and proportions of all occurrences of all terms in the data, that is, term tokens. There are 331 codified term tokens, or 70% of all tokens are codified as unambiguously Welsh by all codification texts that have an entry for the concept. This figure is 10% higher than the proportion of term types in the sample. As was the case with the media data, the differences in the proportions of term types and tokens of codified terms relates to the frequency of occurrence, and the higher proportion of term tokens points to a higher frequency of each term type. On average, there are 3.1 individual occurrences of each codified term type in the sample¹⁵⁰, compared with an average for all codification categories together of 2.6. This is not a very big difference, although, as we shall see, there is a bigger difference between the

¹⁵⁰ Calculated by dividing the number of codified term tokens (331) by the number of codified term types (108).

average number of codified tokens per term and that of not-codified tokens (1.8). I am tempted to conclude that this observation suggests that words that are more central to the organisation are repeated more frequently, and tend to be given institutionally vetted labels. However, looking at the list of terms in Appendix 5, we see that there are a large number of codified terms that occur infrequently. Forty-four occur only once, 27 occur twice, and 9 occur three times (together making up almost $\frac{3}{4}$ of codified terms). The higher occurrence of term tokens in this category seems to be bolstered by a small number of items, most notably *cegin* ('kitchen'), which occurs 31 times in total, more than twice as often as the next most frequent codified term, *stad* ('estate', 14 times). The design of a kitchen happens to be a topic that is discussed in some detail in the *Minafon* recording. Similarly, *stad* occurs frequently because it is the main topic of conversation in *Estates*. Otherwise, the three next most frequent items are central to the work of the Association: *tŷ* ('house', 12 occurrences), *gosod* ('to let', 10 occurrences), and *cwmni* ('company', 10 occurrences).

To turn to the next category, there are 45 term types that are not codified as Welsh by the texts consulted, which accounts for 25% of the term types in the sample. All but one of the terms that fall within the 'not codified' category can be described as borrowings¹⁵¹. Moreover, the terms in the 'not codified' category represent *all* types of lexical borrowing typically used by Welsh speakers, as I discussed in Chapter 4. These include simple items, consisting of one root morpheme, such as *access* and *latch*; more complex terms, formed by English word-formation rules, such as derivatives (e.g. *specification*) and compound words (*reference number*). There are also formulaic expressions (e.g. *health and safety*), as well as complex forms that have gone through a process of ellipsis (*knock-on* for *knock-on effect*). Others still have gone through some process of morphological integration into Welsh, including English verb stems

¹⁵¹ The single term in this category that is not contact-induced, *ailddylunio* ('to redesign', or literally 'to design a second time'), is an anomaly. I assume that it is an oversight that *Geiriadur yr Academi* accepts only *ailgynllunio* for 'redesign', considering that it accepts both *dylunio* and *cynllunio* for 'design'.

given a Welsh verbal derivational suffix (e.g. *supply-io*), and even compound English nouns converted into verbs and put through the same process (e.g. *shortlist-io*). There are loan translations, where the English elements are substituted by Welsh translation equivalents. Of these, some are noun-phrases (e.g. *tâl gwasanaeth* > ‘service charge’), but most are phrasal verbs (e.g. *rhoi ymlaen* > ‘put forward’). In some loan translations, an integrated loan is used as the Welsh translation equivalent of an English verb (e.g. *sortio allan* > ‘sort out’).

As in the case of the codified category, most of the items in this category are not particularly specialist. That is, they are most likely known, as concepts, to participants from outside of their working lives: *access*, *decorator*, *email*, *incident*, *issue*, *pressure*, *symud i mewn* (‘to move in’). What might be said is that they are mostly words that do not denote very central objects within the institution (with the exception of *suspend* – which I discuss above, and possibly *shortlistio* (‘to shortlist’)), and have therefore not been given institutionally vetted designations.

Turning again to term tokens, we find that the percentage of not-codified term tokens (17%, 81 tokens) is lower than that of term types (25%, 45 tokens). Again, this difference relates to frequency of occurrence, and suggests that terms in the ‘not codified’ category are typically used less frequently than those in the codified category (although, again, there isn’t a big difference between the ratio of not-codified terms (1.8) compared with the sample average (2.6)). One of the reasons why words in this category might occur less frequently is that there are alternative designations in the data for a number of the concepts they denote (such as *kitchen*, *email*, and *supplyio* (‘to supply’)). Typically, however, these concepts don’t come up very often in the data (with the notable exception of *kitchen* (which is also denoted by *cegin*, see above) and *issue* (which occurs 15 times in the sample)). We can postulate, then, that these items occur less frequently because they are less central to workplace interaction, and therefore also less likely to be given institutionally vetted designations.

The remaining categories share between them a relatively low proportion of the term types (15% together). Of these, the partly codified and not applicable categories are very small, with only 9 term types between them, or together making up only 5% of the sample. Most of the partly codified items are ambiguous in terms of their semantic scope – in some uses in the data, for example, *gwydr* ('glass') seems to refer to 'window', and it is unclear whether this meaning is included within the scope of *gwydr* in the codification texts.

The not applicable category (four types) is very small (especially in comparison to the same category for the *Syrjeri* data – 16 types). It suggests that most of the work-related vocabulary found in the data is dealt with by lexical planners. Note, however, that I did not include as terms a number of multiword items, that might be considered to designate single objects, since I considered that they were so specific that they would not likely be included in codification texts. These include, for example, *polisi gosod lleol* ('local letting policy'), *rhestr aros gyffredin* ('common waiting list'), *llythyr datgan pryder* ('letter stating concern'), *pwyntia digartra* ('homelessness points'), *tim ardal* ('area teams'). It is also unclear that these should be considered terms at all. I have, however, included multi-word items that I have considered have quite a wide cultural presence, and hence fairly likely to be dealt with by codification texts (*iechyd a diogelwch/health and safety*, *asesiad risg* ('risk assessment'), *date of birth*).

Finally, the mixed category (18 types, or 10%) comprises terms over which the reference texts disagree. We find here a number of borrowed words, such as *cooker*, *insiwrans* ('insurance'), *job*, *opsiwn* ('option'), *plan*, and *printio* ('to print'). In these cases, codification texts tend to disagree whether a borrowing is acceptable, or whether a 'native' word preferable. Typically, the more standardising texts (works of terminology rather than lexicography) tend to prefer native forms, whereas other texts, notably *Geiriadur yr Academi*, are more likely to list a number of different alternatives for a concept, both borrowings and native forms, giving, for example, *cwcer* ('cooker'), *popty* and *ffwrn* as

semantic alternatives. It is in the mixed category that we find the only example of a work-related word from the sample that might be defined as regional – *pres* ('money'). Interestingly, *arian*, a semantic alternative to *pres* is found in the data, but only in derivatives (*ariannol* ('financial'), *ariannu* ('to finance')) or in multiword items (*arian wrth gefn* ('money in reserve')). Considering that this is the only case of regional variation in the sample, it seems clear that the most prevalent form of lexical variation is contact-induced.

Table 6.2 considers whether the general picture varies depending on the recording¹⁵². As I explained above, the recordings differ not only in their subject matter, but in their composition. Some recordings, *Minafon* in particular, include participants that are highly ranked within the organisational hierarchy and who hold a high degree of executive power. Other recordings, *Estates* in particular, include participants of lower rank in the hierarchy. Looking at variation between recordings in some ways helps to look at variation between participants, depending on their hierarchical ranking. But it also allows us to consider the differentiated contexts that might be invoked or construed, depending on the differentiated goals and composition of different speech events. Although the recordings are all meetings conducted within the context of the Association, it is not enough to assume that all meeting language is the same, and does not vary according to participant types, goals and priorities, and the degree to which interaction in a meeting is focused on institutional priorities.

The division of the data into four subsets has resulted in very small numbers for all of the cells in the mixed, partly codified and not applicable categories. I therefore concentrate here only on codified and not-codified terms.

That there is some differentiation between the recordings seems to be borne out in Table 6.2. There is a difference of 15% between the recording with the highest proportion of codified terms (*Voids* with 70%) and the recording with the lowest

¹⁵² I do not present *token* data differentiated by participant group.

proportion (*Publications* with 55%). Similarly, there is a difference of 10% between the recording with the highest proportion of not-codified terms (*Estates* with 30%) and the recordings with the lowest proportion (both *Minafon* and *Voids* have 20%). Two of the recordings seem to group together: the *Voids* and *Minafon* recordings show similar tendencies to each other, in terms of codified and not-codified terms, both showing high levels of codified terms (69% and 67% respectively) and relatively low levels of not-codified terms (both 20%). Similarly, *Publications* and *Estates* show very close percentages of codified terms (55% and 57% respectively), which are substantially lower than those of the other two recordings. However, whereas the percentage of not-codified terms in *Estates* (30%) is substantially higher than in *Minafon* and *Voids*, that of *Publications* is only marginally higher (24%). In fact, what we see in *Publications* is that there are more terms in the ambiguous categories. What we see, then, is that the vocabulary in *Voids* and *Minafon* emerges as highly codified (relatively speaking), that in *Estates* least codified, and *Publications* emerges as taking the middle ground.

	Voids		Minafon		Publications		Estates	
Category	Types	%	Types	%	Types	%	Types	%
Codified	34	69.39	40	66.67	28	54.90	34	56.67
Not codified	10	20.41	12	20.00	12	23.53	18	30.00
Partly cod.	1	2.04	2	3.33	3	5.88	0	0.00
Mixed	3	6.12	6	10.00	7	13.73	6	10.00
n/a	1	2.04	0	0.00	1	1.96	2	3.33
Total terms	49	100	60	100	51	100	60	100

Table 6.2: Housing Association term types according to recording and degree of codification

We can postulate that the differentiated distributional pattern of codified terms between meetings pertains in some way to status level of participants within the organisational hierarchy. The *Minafon* recording, for example, includes primarily higher-status staff members. This recording contrasts quite starkly in terms of its constitution and term categories with the *Estates* meeting, which includes lower-ranking staff members who undertake the quite practical tasks that are directly

tenant- and property-related (dealing with complaints, maintenance work, etc.). What I am describing is an apparent correlation, of course, and not an explanation (see Chapter 3); a meeting with higher-ranking staff members shows a higher proportion of codified terms than a meeting with lower-ranking staff members.

Why might higher-ranking staff use more codified terms than lower ranking staff? We can postulate that ‘correct’ Welsh – defined as codified or perhaps ‘pure’ Welsh – is highly valued within the institution, and that staff members who are more highly-ranked are more likely than less highly-ranked staff members to be concerned with performing identities that have high symbolic value within the institution. Higher ranked participants have perhaps even achieved their higher status as a result of their language practices. Language is often a selection criterion in the recruitment process (and according to Bourdieu’s theory, it is the link between language education and the labour market that consolidates standardisation – see Chapter 2). As I noted in Chapter 4, it is not clear from the Association’s recruitment policy whether language is a formal selection criterion, although it very likely is, to some extent, below the level of consciousness if not formally. Moreover, we do know that the Association offers Welsh-language training and provides lexical codification texts, which suggests that planned language is valued institutionally. Recall also that in Chapter 4 I suggested that class or status group seemed to be implicated in contrasts between ‘standard’ (or ‘posh’) Welsh language practices and more vernacular ways of speaking, a contrast that might be reflected in the apparently differentiated practices of the *Minafon* and *Estates* meeting.

Whilst the *Estates* and *Minafon* recordings provide quite a clear contrast in terms of their constitution and codified terms, the other meetings are more difficult to interpret. Why would the *Voids* meeting, which is made up of staff members of varying status, return a higher proportion of codified terms than *Minafon*? However, it is the constitution of the sample, not of the recording, that is of

interest, and that must account for the proportion of talk that can be attributed to each speaker. 50% of the *Voids* sample (measured as a proportion of the total sample word count) is made up of the talk of just two participants, Nell and Phoebe, both of whom are fairly high ranking within the organisation. What seems to be the case here, then, is that the sample is skewed in their favour. This highlights a methodological problem, where samples are not representative of the wider data. Where the constitution of meetings is fairly uniform, this problem doesn't arise, hence the apparently clear differentiation between *Minafon* and *Estates*.

Conclusions

What can we conclude from this analysis? We have seen that the majority of terms used in the data are codified. However, it is difficult to tell whether this ought to be considered a high proportion or not. One of my research questions on page 45 relates to *whether* speakers used planned vocabulary. 'To a certain extent' is the best we can do here, and it is possibly the best that any analysis can do. The answer must be qualified, firstly because there are a substantial proportion of terms (40%) that are not clearly planned. That is, they are found in the other categories. Secondly, even in the 'codified' category, it is unclear that we ought to consider all these items examples of planned language. For example, can we reasonably conclude that speakers use planned language on the basis that they use the word *tir* ('land'), a word whose date of first attestation in *GPC* goes back to the ninth century? And yet, is it not significant that this form is used, rather than any other, in particular rather than *land*? The use of *tir*, and other very basic vocabulary, is at least in keeping with the purist and standard language ideologies, even if they are not evidence, per se, of the take-up of specific planned forms.

These figures do highlight that the majority of the vocabulary is clearly codified. However, the use of Welsh/codified vocabulary is not necessarily always at the forefront of participants' minds, with 25% of terms coming within the 'not

codified' category, most of which are borrowings, and many of these clearly marked. This category does not include words that are particularly out of reach or highly specialised. That is, participants' use of these words is not a consequence of Welsh designations being restricted to very specific semantic fields. A number of terms, as I noted, have semantic equivalents in the data. We see, for example, the use of both *cegin* and *kitchen*, *iechyd a diogelwch* and *health and safety*, *heddlu* and *police*. Participants' use of borrowed words in cases such as these seems to be a consequence of the intense language contact in Wales and the functional differentiation of Welsh and English. Considering that the practices of the Housing Association, in conducting its business internally through the medium of Welsh, challenges this functional differentiation, it is not surprising that fairly general vocabulary, although in many ways work-related, is to some extent borrowed from English. What these words mean in interaction, however, is not clear from the data. Whether these borrowed words are used because Welsh designations are considered inaccessible, for example, or overly politicised, it is not clear. There is, however, some evidence of contextual differentiation when we look at the distributional patterns of terms according to codification category and recording. Although not-codified terms are used in all recordings, they seem to be more prevalent in some meetings than others. We can postulate that this differentiation may be related to the status level of participants in the organisational hierarchy, although considering the methodological difficulties, the interpretation needs a more nuanced approach. Nonetheless, in none of the recordings does the use of not-codified terms or borrowings seem to contravene institutional norms.

The more ambiguous categories, however, point to the difficulty of defining participants' choice of lexis according to whether it is codified by lexical planners, much in the same way as in the case of the media data. Partly codified terms point to the potential difficulty of interpreting codification texts on the part of users, and of stipulating, on the part of planners, exactly how language ought to be used. The mixed category points to conflicts between codification texts (of

which participants may or may not be aware), in particular where planners disagree as to whether a borrowing should be codified or not. To some extent, however, this is caused by a different approach to planning and presenting codification texts. Texts produced by term planners typically seek to stipulate only one term for each concept, whereas lexicographical works tend to list the words the authors consider are legitimate designations for a particular concept (albeit potentially contextually differentiated), thereby often recording a number of words. As a result, users of these texts may have to use their own judgement as to the suitability of a particular designation for their purposes. As previously noted, it is not clear that text users are aware of the differences between different types of codification text (terminological and lexicographic). Finally, the not applicable category points to the difficulty of providing for all potential language use, since there will always be cases that are not planned for by texts. These largely relate to multiword items, however. It is notable that these more ambiguous categories together make up only 15% of all term types (and 13% of term tokens). Together they do not constitute a very large proportion, and most of these are taken up by the mixed category. It seems, then, that there is not much ambiguity presented by codification texts. However, whether or not participants feel this, it is not clear.

6.4 IN SITU LANGUAGE PRACTICE AT THE HOUSING ASSOCIATION

One of the main questions that need addressing from the previous analysis is whether 60% codified terms (or 70% codified tokens) suggests that lexical planning is implemented in the Association. Conversely, what is the significance of 25% not-codified terms (17% tokens)? One of the problems, as I have noted, in addressing these questions is that a quantitative, decontextualised analysis gives us no indication of the social meaning of these terms in interaction. Do not-codified terms (mostly borrowings) pass as normal in interaction? Or rather, do they pass as normal in institutional talk? That is, are not-codified terms and borrowings used in orderly institutional discourse? Where they are used, do they signal discord? Or are they used because Welsh designations are considered

inaccessible or overly politicised?

It is apparent, from looking at the interactional use of not-codified terms and borrowings (often marked borrowings), that not-codified terms and marked borrowings are indeed used in orderly interaction in the pursuit of institutional goals, although most of the vocabulary can be described as ‘Welsh’, and the vast majority of the syntactic structure. Extract 1, taken from *Voids*, shows the use of borrowings and markedly Welsh lexical items side-by-side. Roughly, the participants are discussing in this extract a problem relating to recording in their internal system how long a new property has been vacant before it is let¹⁵³ – a state of affairs the Association wants to see minimised. Phoebe is highlighting that a fault in this system of recording reflects badly on staff performance, unjustly.

Extract 1

1. Phoebe: ond ma [?] hwn y tri diwrnod bob un (.) yn (property name) (.) os
but this is the three days each (.) at (property name) (.) if
2. di'r **diwrnod cwblhad ymarferol** ar y (.) dydd Gwenar wel da-
the practical [?] completion day is on the (.) Friday well w-
3. (.) dan ni'n cael (.) foidio ond dan ni'm yn (.) dan ni'm gallu osod o
(.) we can (.) void but we can't (.) we can't let it
4. tan y dydd Llun (.) o ran **mesur perfformiad** dio'm yn dangos
until the Monday (.) in terms of measuring performance it doesn't show
5. //xxx\
6. Beth: /dim **guaran****tee** os y bora dydd Llun chwaith (.) os dan ni'm yn
no guarantee if the Monday morning either (.) if we don't
7. cael o tan pump (.) a sa neb yma (.) i bigo fo fyny (.) tan dan ni

¹⁵³ The *Voids* recording can be quite difficult for an outsider to follow, since participants rely quite heavily on background knowledge for referential meaning, as well as on the information in the computer database that is projected onto the screen and structures the discussion. As we see in lines 12-14 here, participants themselves sometimes get confused as to what exactly is the focus of the discussion.

- get it until five (.) and there's nobody here to pick it up (.) until we*
8. *dod i fewn dydd Llun (.) a wedyn ffonio'r pobol sydd fod i gael y
come in on Monday (.) and then phone the people who are supposed
to get the*
9. *lle lle dw mynd [?] i //gael gafael [?] ynyn nhw\
place where am I going to get hold of them?*
10. Phoebe: */xxx\\ (.) ia (.) //ond efo'r\
xxx (.) yes (.) but with the*
11. Beth: */xxx*
12. Phoebe: *to- s- tai newydd //<ydy hein i gyd> [?]\
to- s- these are all new houses*
13. Nell: */tai newydd\
new houses*
14. Beth: *ah dan ni'n dal ar tai newydd
ah we're still on new houses*
15. Phoebe: *efo tai newydd yn (property name)
with new houses at (property name)*
16. Beth: *ia
yes*
17. Phoebe: *oeddau nhw mewn (.) oeddau ni'm yn gallu gosod nhw tan dydd
they were in (.) we couldn't let them until*
18. *Llun nag oeddau
Monday could we*
19. Beth: *nac oeddau
no*
20. Phoebe: *dyddiad //(.) xxx\\ (.) bod nhw'n foid (.) di mynd tri diwrnod cynt
date (.) xxx (.) that they're void (.) has passed three days previously*
21. Beth: */cofio\
I remember*
22. Phoebe: *sy //sy ddim\\ yn (.) adlewyrchu'r **perfformiad** nadi
which doesn't (.) reflect the performance does it*

23. Beth: /do\\
yes
24. (.)
25. Nell: so unwaith eto hwyrach fydd raid ni edrych ar neud **manual** (.)
so once again perhaps we'll have to look at doing a manual (.)
26. **adjustment** ar gyfar hynny (.) um Danni
adjustment for that (.) um Danni

Two borrowings are used in this extract, one by Beth (*guarantee*, line 6) and the other by Nell (*manual adjustment*, line 26). These are both quite marked borrowings, the second in particular, due to its English syntax and the use of the derived noun, *adjustment*. Neither of these borrowings are challenged or corrected here. They are both allowed to pass as normal in the interaction. Moreover, neither of these borrowings seems to be used to change the key of the utterances (see the previous chapter on key). They seem to be part of normal institutional interaction here.

In the case of Beth's *guarantee*, it is arguably used in the defence of members of staff, more than in the pursuit of institutional goals (although these are not necessarily contradictory). That is, Beth's argument in line 6 intensifies Phoebe's argument that a fault in the system reflects badly on staff (a defence of staff within the system). But there is nothing particularly disorderly about her talk here (other than her having misunderstood the context of talk). Nell's *manual adjustment* can be inferred quite unambiguously as being used in the pursuit of institutional goals, to ensure that the organisation's system of management works smoothly and fairly. Nell even ensures that her requirement for a manual adjustment gets recorded in the official minutes of the meeting (this is what she means in line 26 with "um (.) Danni", since Danni is the one who is taking the minutes). Her use of *manual adjustment* here, then, is accepted as normal, or appropriate within the interaction, in talk that is focused on institutional priorities and that is to be noted in the official record of the meeting.

What we see in this extract, then, is the use of marked borrowings and not-codified terms in talk that is presented as, and seems to be accepted as, normal. Moreover, borrowings are used by two participants that are positioned on quite different ranks in the organisational hierarchy – Beth nearer the bottom and Nell nearer the top. The same proviso applies here, of course, as applied in the previous chapter, that it is not clear how participants perceive the language of their co-participants, or perceive their co-participants on the basis of the language they use. For example, Beth’s spoken style strikes me as quite odd. She speaks quite slowly, and she doesn’t mutate where I would expect her to according to my own sense of what is normal (in this Extract 2: *tan pump* rather than *tan bump*, and *ffonio’r pobol* rather than *ffonio’r bobol*). It is possible that her co-participants also judge her the same way. But my focus here is on what is allowed to come into the interaction, that is, what is allowed to pass as normal.

Alongside these borrowings, however, we also see some quite markedly Welsh lexical items, mostly used by Phoebe: *diwrnod cwblhad ymarferol* (‘practical completion date’) and *(mesur) perfformiad* (‘(measuring) performance’). I call these lexical items, since they appear to me to be objects of knowledge within the Association, which are used in the running, management and monitoring of its work. ‘Performance’, for example, refers to how well a member of staff within an organisation carries out their role, according to certain evaluative measures. ‘Performance’ has become a highly salient term within the management of business and other organisations. Is there any difference between the use of these markedly Welsh items and the borrowings discussed from the extract? There doesn’t seem to be, in terms of the overall style of the interaction, or the intention of the participants. Is there any reason why these markedly Welsh items are used and not borrowings?

I argue here that, whilst all the lexical content in the extract is allowed to come into the discourse, it is possible that these markedly Welsh lexical items are made up of Welsh morphological content because they are more salient within the

organisation of the Association's business, part of the community of practice's shared repertoire of resources – such as *mesur perfformiad*. Their salience within the organisation, as semi-officialised, I assume, leads to their having been given Welsh-language designations. I made a similar argument above, suggesting that concepts more salient within the organisation's structure are given institutionally vetted labels. By contrast, I argue that a *manual adjustment* is not an intrinsic part of a system. It is the 'human touch' that is used where the system does not function maximally, a concept denoted metaphorically by the image of a 'hand' (*manual* itself being a borrowing from Old French (Oxford University Press, 2013)). Nell probably doesn't have easy access to a 'Welsh' equivalent, because *manual adjustment* isn't an object of knowledge that is salient in the Association, and hence there isn't a Welsh language designation that Nell can deploy here (apart from the fact that *manual adjustment* is difficult to formulate in Welsh).

Note, however, that there *are* other ways Nell could make her reference, but probably not using the specific object of *manual adjustment*. Rather, she would probably have to paraphrase by referring to an action: e.g. *bydd rhaid ni edrych ar newid hynna ein hunain* ('we'll have to look into changing that ourselves'). The point here is that Nell doesn't use a different formulation, since she doesn't need to: *manual adjustment* is accessible, is concise and is perfectly appropriate within this particular interaction. What this extract suggests, then, is that salient, frequently-repeated vocabulary might likely be more 'Welsh', but that within the context of the interaction, purism isn't naturalised enough to warrant 'Welsh' vocabulary for concepts of lower frequency, that might be difficult to access, on the part of the speaker and the recipient.

However, it is not clear that this is the case all of the time. I have so far presented a very small number of words in one small part of one of the recordings. Looking at other parts of the data, we see that there are cases where borrowing seems to be presented as not appropriate. In Extract 2, taken from the *Minafon* recording,

we see Gaby correcting her own talk, which I argue is motivated by purism. In the wider episode from which this extract has been taken, participants are discussing how they should categorise, for financial purposes, the weekly work of testing fire alarms in the flats at the *Minafon* development. In Extract 2, Gaby is presenting the broad theme to be discussed, that of testing the fire alarms. She then explains in some detail what in particular she wants the participants to focus their attention on: whether the work of testing fire alarms should be paid for out of tenants' rents or out of their service charge payments.

Extract 2

1. Phil: ok (.) oes na rwbeth arall am rhenti (.) ta- neu tâl gwasanaeth
ok (.) is there anything else about rents (.) or service charge
2. dan ni'n angen <delio gyda> [?] bore ma
we need to <deal with> [?] this morning?
3. (2.0)
4. Gaby: //oes:\
yes
5. Nell: /xxx tâl gwasanaeth\\ xxx
xxx service charge xxx
6. Gaby: ma na (.) yr issues ma (.) sydd yn ffitio fewn i radda efo'r
there are (.) these issues (.) that fit in to some extent with this
7. gofalydd campws ma nes mlaen o ran (.) be sy'n cael ei gynnwys
campus caretaker later on in terms of (.) what's included
8. yn y rent a be sy'n cael ei gynnwys yn tâl gwasanaeth (.) sef yr
in the rent and what's included in the service charge (.) that is
9. elfenna (.) larwm tân **profi**r larwm tân (.) petha fel anti-scal- scal-
the (.) fire alarm elements testing the fire alarm (.) things like anti-scal- scal-
10. (.) ffordd dwi'n ddeud o Karl [?] =
(.) how do I say it, Karl? [?] =
11. Nell: = scald [?]

12. Karl: //<scald valves\> [?]
13. Gaby: /<xxx\ \ scald> [laughs] (.) valves (.) um (.) petha fela (.) ar y
<xxx scald> [laughs] (.) valves (.) um (.) things like that (.) at the
14. funud dyn nhw ddim (.) dyn nhw ddim yn y rent: (.) dan ni'n t- dan
moment they're not (.) they're not in the rent (.) we're t- we're
15. ni'n codi tâl gwasanaeth os di rywun yn byw mewn fflat (.) i rywun
charging service charge if somebody lives in a flat (.) for somebody
16. mynd yna i **brofi**'r larwm tân (.) ma na ddisgwyl oherwydd
to go there to test the fire alarm (.) it's expected because of
17. strwythur (.) strategaeth tân y cynllun yma (.) bod raid i'r laryma
the (.) fire strategy structure of this scheme (.) that the fire
18. tân (.) a'r (.) gola (.) diogelwch (.) argyfwng xxx diogelwch (.) i
alarms (.) and the (.) safety (.) light (.) emergency xxx safety (.)
19. gyd gael ei **hecio**'n wythnosol (.) so ma na oblygiada cost o
all have to be checked weekly (.) so there are cost obligations for
20. hynna dyna pam i radda dan ni di trio cael y swydd gofalydd ma
that that's why to some extent we've tried to get this caretaker job
21. (.) mewn lle (.) ond cwestiynu dwi (.) ar y funud (.) ma (.) ma'r
(.) in place (.) but I'm questioning (.) at the moment (.) the (.) the
22. costa (.) i **hecio**'r (.) **profi**'r larwm tân i gyd yn y tâl gwasanaeth
costs (.) to check the (.) test the fire alarm are all in the service charge
23. ydy o'n xx disgwyledig (.) bod y tenantiaid Minafon oherwydd
is it xx expected (.) that the Minafon tenants because of
24. strwythur yr adeilad (.) yn gorfod talu o'u //tâl gwasanaeth\
the structure of the building (.) have to pay out of their service
charge?
25. Phil: /ma [?] **profi**\ (.) larwm tân yn ddyletswydd landlord (.) felly
testing (.) a fire alarm is a landlord's duty (.) so
26. allan o rent <dyle hwn ddod> [?]
<this should come> [?] out of rent

To a large extent, this extract is very similar to the previous: there are a number

of fairly salient workplace words that are saliently Welsh, such as *tâl gwasanaeth* ('service charge'), *gofalydd campws* ('campus caretaker', salient within the context of the *Minafon* project group), *larwm tân* ('fire alarm'), as well as some borrowings that are either not salient within the work of the Housing Association (*anti-scald valves*) or difficult to translate into Welsh (*issue* – a word that is very commonly used, but for which there is no exact and equally useful Welsh semantic translation). However, one difference is Gaby's use of the words *profi* ('to test') and *checio* ('to check'). Gaby uses these verbs to refer to the action that is to be carried out on the fire alarms. First she chooses the verb *profi* (literally, 'to test') in line 9. She repeats this formulation in line 16. However, in line 19, in a passive construction, she uses a different expression – the verb *checio* (borrowed and adapted from English *to check*) – to refer to the same action. Gaby repeats *checio* in line 22. However, here she immediately reformulates the reference again, switching back to *profi*.

Why does Gaby make this switch? She does not seem to be presenting two lexical items purposefully as semantic equivalents here. She does not, for example, seem to be using the same kind of purposeful reformulating apposition that we saw in the *Syrjeri* analysis. This is suggested by her hesitation before *profi* as well as the delayed onset of the initial /p/. She seems to have halted the flow of her talk, in order to erase her prior use of *checio* and replace it with *profi*. That is, her switch sounds very much like self-repair. Her reflexive mechanism seems to have kicked in here, telling her that *checio* isn't appropriate (whereas she doesn't seem to have considered this in her previous use of *checio* in line 19).

Why is *checio* inappropriate here for Gaby's purposes, and what work does *profi* do differently? It is possible that Gaby is orienting to the semantic difference between *profi* and *checio* (my translation of the extract reflects this difference, where I have translated *profi* as *test* and *checio* as *check*). However, there does not seem to be any real need for semantic specificity here, any more than in her other references to the action (in lines 9, 16 and 19). Another possibility is that

Gaby becomes aware of the loan status of *checio* and reformulates for puristic reasons.

I should note that both *checio* and *profi* are, in fact borrowings. *Checio* is a borrowing from English, while *profi* is from Latin (H. Lewis, 1943), related to the English *prove*, but not directly borrowed from it. *Checio* is probably more salient to Gaby as a borrowing than *profi*, because it is more obviously equivalent to another form in her linguistic repertoire (*to check*). Moreover, the phonological and morphological shape of *checio* make it salient as a borrowing: the stigmatised /tʃ/ phoneme (it was not even assigned a grapheme in the Welsh alphabet)¹⁵⁴, and the root morpheme *check* plus verbal suffix *-(i)o*. The latter is a highly productive strategy for the incorporation of English verbs into Welsh (Stammers, 2009).

We might conclude, therefore, that Gaby's switch from *checio* to *profi* is manifesting a purist ideology in her showing preference for *profi* over *checio* here, at least as far as her linguistic repertoire allows her to, and in as far as she feels she needs to. By these qualifications, I'm drawing attention to the fact that she doesn't show preference for 'native' forms all the time, as is evidenced by her use of other borrowings in the extract: *anti-scald valves*, *issues*, *rent*, *fflat*, *cost*, *ffitio* ('to fit') and *tenant*. It may be that some of these forms aren't salient as borrowings. Some of them are, in fact, accepted by lexical planners – *rent*, *fflat*, *cost*, *ffitio*. On the other hand, in the case of *anti-scald valve*, although markedly a borrowing, it is clearly an item that is not firmly established in her linguistic repertoire, as we see by her searching for the form in lines 9-13. Arguably, within the context of this interaction, Gaby doesn't consider that the use of the form *anti-scald valve* is inappropriate. Despite Gaby's self-repair in line 22, purism isn't naturalised enough in this episode for Gaby to make the effort to search for a suitable Welsh semantic equivalent for *anti-scald valve* that would require innovation on Gaby's

¹⁵⁴ I have used <ch> to represent the /tʃ/ in *checio*, despite <ch> usually representing /χ/ in Welsh orthography.

part and effort on the part of the recipients.

We might leave the analysis there, happy that Gaby is manifesting a purist ideology within certain limits. But my analysis so far hasn't answered why she does this. It should be noted that Gaby shows a great deal of reflexivity throughout the recording. She seems to choose her words quite carefully, using a lot of self-initiated repairs as well as metalinguistic commentary, more so than the other participants. Gaby is under a great deal of pressure in this meeting. This might explain her heightened metapragmatic awareness and precision in her choice of linguistic forms. She is the participant with the most responsibility for presenting information for discussion in this almost three-hour long meeting, and who has prepared much of the ample documentation. The meeting is held in order to ensure the smooth running of the development of the *Minafon* flats. The success of the development rests to a great extent on Gaby's coming to the meeting well-prepared, and on her performance in the meeting. She must be feeling some pressure. Moreover, in Extract 2, as with much of the recording, there is a semi-dialogue structure between Phil and Gaby. Phil typically asks questions, Gaby gives some information and Phil offers his opinion or a reformulation of what Gaby has said. The significance of this semi-dialogue structure is that as well as chairing the meeting, Phil is ranked very highly in the organisational hierarchy. Therefore, Gaby is under double the pressure to perform competently. Part of that performance is the language she uses in her talk. It is possible, then, that Gaby's self-repair is motivated by a wish to use linguistic resources she perceives to be highly valued within the institution. It is not necessarily the case, then, that she implements purism here due to a personal commitment to certain kind of Welsh language, but because she perceives that such language is institutionally legitimated within the Association (although both interpretations may be valid at the same time).

This interpretation also helps explain the variation we saw in the previous section in the proportion of codified and not-codified terms found between

meetings. One of the meetings that had a higher proportion of codified and lower proportion of not-codified terms was *Minafon*, the meeting from which Extract 2 has been taken. It is possible that the presence of highly ranked staff in this meeting has an effect on the language practices and metalinguistic awareness of participants, motivating them to use what they perceive to be legitimate language within the institution. Indeed, Phil himself uses such language. Whereas he does use borrowings, even marked borrowings sometimes, he doesn't seem to use very many. He only uses *specification*, *job*, *kerb*, *equity*, *issue*, *utilities* and *knock-on (effect)*. Other than *job* (and possibly *specification*), all these words are difficult to translate into Welsh.

Here is some evidence, then, that purist ideology may be highly valued within the institution. Nonetheless, this assertion must be qualified by the point already made that purism isn't entirely naturalised. In fact, I should emphasise that there aren't many examples in the data that echo Gaby's self-repair. I have found only two other examples of self-repair where the speaker potentially switches for a word from another language. One of these is Gaby's own, where she corrects *bloc o flats* ('block of flats') with *bloc o fflatia*, changing the 'English' plural suffix -s for the 'Welsh' -ia. In the other example, by Pam in *Voids*, the switch seems to be from Welsh to English: "paid ti â m- (.) derio mynd i weld o xxx medda fo" ("don't you m- (.) dare go see him xxx" he said'). It is uncertain, but quite possible, that Pam's "m-" before the pause is the initial phoneme of a truncated *meiddio* ('to dare'), the semantic equivalent of *derio* (English *dare* + -io), which she articulates fully. If so, it would mean that she repairs Welsh with English, or rather, with an integrated English borrowing. Pam is quite a lively performer, whose spoken style is typically vivid and peppered with direct speech that help bring alive her narrative. I suggest that the switch from *meiddio* to *derio* may have been motivated by the more dramatic effect that the initial plosive consonant /d/ of the latter offered over the nasal /m/ of the former. Together with an accented first syllable and a short pause, her articulation of *derio* emphasises the intensity of the message that the quoted speaker was relaying to her. Clearly this switch is

not motivated by purism, and it may be that purism is completely absent from, and in conflict with, her dramatic spoken style, since it would restrict her ability to colour her talk in this way. But it is worth emphasising that, within this interaction, this kind of talk is allowed, although not everybody in the interaction does the same as Pam.

It is worth pausing here to consider the difference between Pam's and Gaby's self-repair examples, which come from different meetings. Although there seem to be differences in their personal spoken styles (Pam is typically vivid and dramatic, whilst Gaby seems to be more self-censoring¹⁵⁵), I suggest that the difference is also an effect of the different contexts of interaction, including what is said, the overall aims of the meetings, Pam's and Gaby's positions within the organisation, and their co-participants. I have highlighted how Gaby is under some pressure to perform in *Minafon*. The stakes are quite high for her. In fact, the high stakes are quite prevalent in this meeting, not only for Gaby but for other participants too. This is manifested best, perhaps, in participants typically showing some deference to highly-ranked Phil, who takes on the role of chairing. In the *Voids* meeting, on the other hand, participants are clearly under less pressure to perform and please. Although Nell is highly ranked, is their boss and chairs the meeting, it is much more clearly a meeting between equals. Moreover, Nell shows some care for her staff, noting some concern and committing to act where participants complain about some working practices that potentially impact on their personal safety.

My point here is that Pam's vivid spoken style – and switch to an English borrowing – would probably not be appropriate talk within the context of the *Minafon* meeting. Conversely, Gaby's purist self-censoring would not be needed – and possibly not appropriate – within the context of the *Voids* meeting. Although (marked) borrowings and markedly Welsh words are used in both meetings, it

¹⁵⁵ Although it is difficult to assess the effect of the immediate situation on Gaby's talk, because *Minafon* is the only recording she makes an appearance in.

seems to me that the ideology of purism is more firmly embedded within the *Minafon* meeting than in *Voids*. I suggest that this is a result of the valuing of ‘pure’ Welsh as ‘correct’ Welsh in what is viewed as ‘competent’ performance in the higher levels of the institution.

Nonetheless, if purism is highly valued within the institution, it is not thoroughly naturalised within the spoken interaction – even in *Minafon* – considering that there is such little self-repair (and no other-repair) but that borrowings, including marked ones, are quite ubiquitous. In the Housing Association purism and runs into participants’ other considerations that restrict its applicability. These considerations include in situ demands of the interaction. The following two extracts, viewed together, point to how participants can vary their language use according to their goals. These extracts are both taken from the *Voids* recording. They concern the linguistic resources Nell uses to refer to ‘vouchers’ in a scheme that is being piloted by the Association whereby they give tenants vouchers to buy paint to decorate their properties. This scheme is one of the items on the agenda for the meeting. In Extract 3, which comes from the beginning of the meeting, Nell, who is chairing, is quickly running through the items of the minutes of the previous meeting, listing which of them they are going to discuss that day. The turn takes up an entire minute, but I’ve only presented a part of it here.

Extract 3

1. // (chatting) \
2. Nell: /felly (.) efo (.) hynny o beth (.) um \ \ fysan ni’n cael cychwyn um (.)
so (.) with (.) that (.) um could we start? um (.)
3. mae gynnon ni nodiada cyfarfod mae o mynd yn ôl i fis Gorffennaf
um (.) we’ve got minutes it goes back to July
4. bellach [?] so ma na lot <ono fo> [?] reit hanesyddol (.) so dwi dwi’m
now [?] so a lot <of it> [?] is quite historical (.) so I- I don’t
5. isho mynd trwyddo fo (.) yn fanwl oni bai bod gynnoch chi ryw (.)

want to go through it (.) in detail unless you've got some (.)

6. faterion penodol dach chi isho bigo fyny (.) um (1.5) ti gwbod dan
specific matters you want to pick up (.) um (1.5) you know we're
7. ni'n sôn am y (.) dan ni isho dal i fyny fo'r (.) cynllun **taleba** peintio
talking about the (.) we want to catch up with the (.) painting
vouchers scheme
8. beth bynnag yn does (.) a dan ni mynd trwy'r targeda a (.) timod y
anyway don't we (.) and we're going through the targets and (.) you
know the
9. foidia (.) felly oni bai bod na wbath (.) mawr (.) um (2.5) a wedyn ma
voids (.) so unless there's anything (.) major (.) um (2.5) and then we've
10. gynnon ni (.) dan ni isho pigo fyny efo [?] eiddo anodd i osod (.) a
got (.) we want to pick up with [?] property that's difficult to let (.) and
11. wedyn wrth gwrs efo'r (.) Lean Systams wedyn
then of course with the (.) Lean Systems then

After a long interval (over an hour), where participants have been discussing other items on the agenda, Nell seeks to negotiate the transition between one topic for discussion and the next, the next being the painting vouchers scheme. This sequence is presented in Extract 4.

Extract 4

1. Nell: um (.) oes a unrhyw beth (.) arall efo [?] hynny cyn i ni mynd ymlaen
um (.) is there anything (.) else with [?] that before we go on?
2. dan ni di trafod am yr eiddo sy'n anodd i osod do wrth fynd yn ein
we've talked about the property that's difficult to let haven't we as
we've gone
3. blaena os a rwbath arall dach chi xxx isho bigo fyny (.) efo hynna
along is there anything else you xxx want to pick up (.) with that?
4. (1.5) jyst isho pigo fyny rili efo (.) um (.) cynllun ef- timod bod ni'n
(1.5) just want to pick up really with (.) um (.) the scheme wi- you
know that we're
5. (.) peilota fo'r och- gwaith (.) peintio (.) efo **vouchers** (.) timod yr

(.) *piloting with the si- painting* (.) *work* (.) *with vouchers* (.) *you know the*

6. e-bost gaethon ni gyn ti'n ddiweddar do Paul ynglŷn â (.)
email we got from you recently didn't we Paul about (.)
7. //ynglŷn â arbedion\
about savings
8. Jim: /mae o di gweithio\\ (.) mae di gweithio mewn rei (.) ond dydy o
it's worked (.) *it's worked in some* (.) *but it's*
9. ddim (.) naeth o'm gweithio i un naddo Pam
not (.) *it didn't work for one did it Pam*

In these extracts, we see Nell using two different lexical items to refer to 'voucher', the 'Welsh' word *taleb* in the first extract, and the borrowing *voucher* in the second. But in considering why Nell's use varies here, I consider the lexical items within the broader referring expressions (Schiffrin, 2006) she uses to refer to the 'painting vouchers scheme'. Both referring expressions are very different, and the variation in Nell's use of *taleb* in one case and *voucher* in the other seems to coincide with this difference.

In Extract 3, Nell uses *taleb* in a compound lexical item, *cynllun taleba peintio* (line 7). This lexical item is a noun phrase, which follows typically Welsh word-formation rules. The noun phrase is made up of a head (the noun *cynllun* ('scheme')) and subsequent modifier (*taleba peintio* ('painting vouchers')). The modifier itself is a noun phrase, which consists of a head (the noun *taleba* ('vouchers')) and subsequent modifier (the infinitive verb *peintio* ('to paint')). This is a very neat and concise formulation to refer to the scheme.

On the other hand, in Extract 4, Nell uses a strikingly different formulation to refer to the scheme. The entire formulation can be seen as consisting of her speech between *cynllun* ('scheme') in line 4, and *vouchers* in line 5, or even *arbedion* ('savings') in line 7. This formulation is much less neat and concise than the one Nell uses in Extract 3. She doesn't condense the reference into a single

lexical item, following word formation rules, like she does with *cynllun taleba peintio*. Rather, following a long filled pause, she uses a long and complex structure, starting with a noun followed by a cut-off preposition and repair (“cynllun ef- timod bod ni’n...” (‘scheme wi- you know that we’re’)). The repair (not a relative clause as the translation might suggest) comprises a new sentence and a sentence-internal repair (“’r och- gwaith (.) peintio” (‘the si- painting (.) work’)). She then elaborates further, referring to an email received on the subject by one of her co-participants.

Why does Nell use such different formulations in making the same reference in these extracts? The contexts are very similar: she is acting in a professional capacity, addressing all participants, she is leading and steering the discussion, making decisions regarding what topic is going to be discussed next. However, there are two points where the extracts differ. Firstly, *cynllun taleba peintio* is not strictly new in Extract 3. It is clear that Nell is referring directly to the written minutes of the previous meeting, the *nodiada cyfarfod* (‘minutes’) she refers to in line 3. She is listing those items in the minutes that they will be discussing during that day’s meeting. Therefore, *cynllun taleba peintio* already has a history; it is a repetition of a formulation already current. It has been primed in the prior text. On the other hand, in Extract 4, Nell is not borrowing her words from the minutes, but selecting more freely from her own lexicogrammatical resources.

Secondly, Nell’s communicative aims differ in both extracts. She seems to be trying to be much more explicit in Extract 4 than in Extract 3. Reference is crucial in the conveyance of meaning, and the initial referring expression is typically most crucial (Schiffrin, 2006), since subsequent referrals always refer back to them:

“the most crucial features of each utterance, the feature which a listener must minimally grasp in order to begin to understand the utterance, is the expression used to identify what the speaker is

talking about” (Brown, 1995: 62 in Schiffrin, 2006: 33).

However, such an approach to discourse can be too focused on the collaborative nature of interaction. A more critical approach would highlight that interactants do not always engage with each other on an equal footing, and do not always seek to be entirely open with each other. In some cases, it is within the speaker’s interests to be explicit, but in others not (deception is an obvious example). Nell is the chair of this meeting (and ranked highly in the organisational hierarchy), and she has considerable power to steer the discussion in the direction she wants to take it. Nell’s being more or less explicit can be explained from this perspective. In Extract 4 her aim seems to be to introduce the topic for discussion in a way that makes it clear to her colleagues what she wants them to focus their minds and their talk on. This requires a much more detailed and descriptive (and thus less neat and concise) formulation. In fact, Nell goes on with her referring expression until her co-participants explicitly indicate that they have understood. Jim does so in line 8.

Conversely, in Extract 3, Nell doesn’t have the same need for clarity and explicitness. In fact, the opposite would be more in her interest, since she states quite clearly (lines 4-5) that she’d rather move on with the rest of the agenda than discuss the minutes of the previous meeting, from which she is referring to the painting vouchers scheme. Note that she does include the condition ‘unless you’ve got some specific issues you want to discuss’, thus opening up a transition-relevance place (Sacks, Schegloff, & Jefferson, 1974) which would allow one of her co-participants to take the floor. However, she doesn’t leave the transition-relevance place open for long, resuming the floor after just 1.5 seconds’ silence.

Nell’s use of the concise formulation, readily available from the minutes, allows her to quickly refer to the scheme and move on. *Cynllun taleba peintio* doesn’t say much to the uninitiated, and may not be totally clear to a member of the

community of practice unless they are given more contextual information. Nell does give more information in Extract 4 where her co-participants' understanding is important to her, because here her co-participants' understanding of what she's referring to is crucial to how she wishes for the meeting to proceed.

Returning to *taleb* and *voucher* specifically, it is notable that in trying to be more explicit in Extract 4, Nell not only uses a different formulation than that in Extract 3, but she also uses a different word to denote the concept 'voucher'. Part of being more explicit and seeking to ensure that she is being understood involves using the word *voucher* rather than *taleb*. *Voucher* seems to be the preferred denotation of the other participants, and Nell is aware of this. In the eleven remaining cases where 'voucher' is referred to in the recording (by six different participants, including Nell), *voucher* is used, not *taleb*. Therefore, we can consider that Nell uses *voucher* to be explicit, in order to engage her colleagues by using their preferred word. This is part of her communicative competence.

What we find in the case of *taleb/voucher*, then, is that the use of the borrowing *voucher* is perfectly appropriate in the spoken interaction in Extract 4, more appropriate than *taleb* would be. It is notable, however, that it is *taleb* that has been used in the written minutes of the meeting and not *voucher*. This, no doubt, is partly motivated by the fact that *voucher* is difficult to write in Welsh orthography, and that using the English spelling *voucher* would introduce a letter that is not recognised as Welsh (the letter *v*). But it is also motivated, I believe, by the more purist norms that usually guide formal writing in Welsh¹⁵⁶.

That *voucher* is appropriate in Extract 4 might be seen to contradict what I

¹⁵⁶ A similar, writing-related, example is found in the use of *kerb*. It is used and presented as normal in the spoken interaction (in *Minafon*), pronounced variously as /kərb/ and /kɜ:b/. But in a written report that was part of the documentation for the meeting, *kerb* is written in inverted commas, highlighting its status as a borrowing and that its spelling does not follow the conventions of Welsh orthography (partly because of the letter *k*, which is not found in the Welsh alphabet).

argued in the case of *checio* in Extract 2 – that Gaby corrects her use of *checio* because she considers it inappropriate, on account of it being a borrowing. It is possible that, as a fairly unfamiliar word, *taleb* is less likely to be considered accessible than *profi*, and therefore speakers are more likely to be tolerant of using the borrowing *voucher*. But there are other interpretations that might be used to explain the difference between Nell’s use of *voucher* and Gaby’s self-correction. Firstly, it is possible that Gaby is more mindful of her use of ‘correct’ language in the first place. She is one of the only participants in the dataset who shows evidence of self-repair that seems to be motivated by purism or language ‘correctness’. Conversely, Nell seems to use more borrowings more freely. Indeed, in the interaction that follows the end of Extract 2, Nell uses *checio* to refer to the action carried out on the fire alarms, where others continue to use *profi*. It is possible, then, that there are personal differences between Nell and Gaby. On the other hand, it is also possible that the difference arises, in part at least, from their different roles within the Association. Nell is higher on the organisational hierarchy than Gaby, suggesting that she may not feel such need to display her ability in using institutionally legitimate language as does Gaby. She is not under such pressure to perform.

It is also important to note that Gaby and Nell are speaking to different audiences in Extracts 2 and 3/4. Whilst Gaby is interacting with higher-level staff in Extract 2, especially Phil, in Extracts 3 and 4, Nell is speaking to lower-level staff. The significance here is that it may be that participants are more likely to orient to an institutionally legitimate language when the stakes are higher. By contrast, it may be that in talk amongst lower-level staff, the use of more vernacular practice, which typically involves more borrowings, may be viewed as more appropriate. This would be supported by my inference that Nell, in Extract 4, uses *voucher* for reasons of accessibility, as well as to use the preferred language of her co-participants.

From this perspective, it is instructive, I think, to consider language practices in

the *Estates* meeting. This meeting emerged as one of the meetings with the lowest proportion of codified terms, and as the meeting with the highest proportion of not-codified terms. I explained this, in part, as being related to the participants' status level within the Association, their linguistic repertoires and the topics of conversation. Can we shed further light on this meeting by considering language use in interaction? Parts of this recording can be said to involve talk where participants speak in a more impassioned, less abstracted way than what is found, for example, in the *Minafon* recording, which primarily focuses on project management, budgeting, etc. The participants in the *Estates* meeting are lower-level staff members, many of whom deal on a day-to-day basis with the Association's tenants. Their perspective, then, relates much more closely to the human element of the Association's work than in the other meetings, in particular in *Voids* and *Minafon*. The Association seeks to implement a system for its work, and lower-level staff members are charged with implementing this system on the ground. At times, these members of staff can see systemic faults or hurdles that have an impact on the lives of the Association's tenants. The meeting itself has been held to discuss a recent survey of their properties – the estate visits – which they had conducted by questioning tenants about the condition of their homes. In particular, some staff members weren't sure how to deal with tenants relating problems at their properties that had previously been reported, but that had not yet been attended to due to budget constraints. These staff members were concerned that this might impact negatively on their own relationships with the tenants, a relationship that was important for them to maintain. At times, this resulted in some frustration, and staff members can voice empathy with their tenants, as we see in Extract 5. In this extract, the participants are discussing one of their tenants. Todd had previously introduced the topic, but had left the room to answer a telephone call. The extract begins on Todd's return.

Extract 5

1. Susan: uh sôn am yr ardd o'n ni de
 uh we were talking about the garden weren't we

2. Todd: na uh y gegin i ddechra arni de //timod oedd (.) tŷ ei hun\ (.)
no uh the kitchen to start with you know the (.) house itself was (.)
3. Susan: /oh y gegin\\
oh the kitchen
4. Todd: xxx doedd chwara //teg doedd\
xxx wasn't it fair play wasn't it
5. Susan: /**immaculate**\\ =
6. Paul: = s- mae na ers blynyddoedd
she's been there for years?
7. Todd: yndy (.) ers (.) ers (.) **seventeen years** =
yes (.) for (.) for (.) seventeen years
8. Jim: //xxx timod\
xxx you know
9. SXX: = /a mae di cael **hell** [laughs] efo\\ (.) xxx //xxx dydy\
and she's had hell with (.) xxx hasn't she
10. Todd: /oedd (.)\\ ond di hi'm yn cael gegin am ddwy flynnadd arall
xxx (.) yes (.) but she's not getting a kitchen for another two years
11. (.) a ma hi yna ers (.) **un-deg-saith mlynadd** (.) ac oedd hi'n
(.) and she's been there for (.) seventeen years and she was
12. deud y (.) tŷ'n is lawr neu'n is lawr wedyn w bath de (.) ers dwi
saying the (.) next house down or the one after that or something right (.) "since I've been
13. ma ma ma hwnna di cael dau **gitchen** meddai (.) so os ti malu dy
here he's had two kitchens" she said (.) "so if you trash your
14. **gegin** ti'n cael lly (.) //<ac oedd hi> [?]\
kitchen you get (one) like" (.) and she was
15. Lauren: /ie nes i\\ ofyn ddoe efo (.) xxx dŷ (place name) nhw'n byw yn y
yes I asked yesterday with (.) xxx house in (place name) them been living in the
16. tai ers **fifteen years** ac oedden nhw'n **due** i gael un (.) flwyddyn
houses for fifteen years and they were due to get one (.) this

17. //yma\
year
18. Paul: /ond [?] **eight****een** di'r cyl- cylch wan de
but the cycle is eighteen now isn't it
19. Lauren: m- ond oeddau nhw'n
but they were
20. Phoebe: **eighteen**
21. Jim: //xxx\
[...]
22. Susan: /argol sôn am dy glân\\ (.) gaethon ni lot o //dai glân yn\
god, talk about a clean house (.) we got a lot of clean houses in
23. Paul: /xxx\
[...]
24. Susan: (place name) [laughs]
[...]
25. Phoebe: sut sa ni dan ni mynd i symud ymlae- efo'r ymweliada stad dan
how are we we're going to move on- with the estate visits we
26. ni isho dal i g- i gynnal nhw ond dan ni isho ella (personal name)
want to carry on h- holding them but we want perhaps (personal name)
27. a (personal name) on board ond a- (.) dan ni dan ni isho (.) gwell
and (personal name) on board but a- (.) we want (.) better
28. atebion a dim (.) gynnon ni mae'r atebion bob tro naci
answer and it's not (.) always us that has them is it

This conversation is one example amongst many in this recording of tenant personal histories, where staff members express empathy with tenants. The meeting itself has no clear structure: it has been convened to discuss how they should proceed with the estate visits, but there is no clear strategy. As we see at the end of this extract, Phoebe, acting within her institutional role as leader of this team, tries to steer the conversation in a strategic direction and towards the pursuit of institutional goals – seeking to find a solution to the problem of knowing what to tell tenants following the estates visits. But Phoebe is thwarted

in her efforts by her co-participants returning to the stories of individual tenants, such as the one seen in Extract 5.

Frames in this meeting shift, from institutional, where talk is often abstract and concerned with strategic action (mostly Phoebe), to a more personal frame, where a concern with common humanity is prioritised. We see this shifting in Phoebe's attempt, in lines 25-28, at steering the talk away from the tenant-focused discussion. Interaction in this meeting, then, is often outside of, or at the peripheries of, an institutional frame.

It is clear that some of the participants feel frustration at the system within which they work. This is not to criticise the Housing Association for its working practices. The Association, a social enterprise, has a very difficult task and is constrained by budgets like any other organisation. It is also clear that it has a caring staff. But the hierarchical organisation through which work is managed necessarily entails that those in lower-ranking posts have less power to change that system, although they *are* the people who work on the ground with tenants and see first-hand where the system runs into difficulties. It is inevitable, I think, that these staff members feel some powerlessness and hence frustration. Meetings such as this, where they are speaking amongst equals, are an opportunity for them to voice their frustrations and to let off steam. Their talk comes across as fairly emotive. Is it possible to say that participants in this meeting voice this frustration and empathy with tenants through linguistic cues, including lexical resources?

It is possible that where the frame shifts from the institutional to the more personal, where participants become more impassioned on behalf of their tenants, they more readily use borrowings and show more signs of syncretic bilingualism. By *syncretic bilingualism*, I mean a practice or state where the bilingual does not seek to keep the linguistic resources of his/her two languages separate, where the opposition between both languages is not construed as

meaningful (Hill, 2000). It is notable, I believe, that participants use a number of borrowings in this extract, in particular English-language numerals *seventeen* (line 7), *fifteen* (line 16) and *eighteen* (lines 18 and 20). Moreover, these are used, with other English borrowings, in syntactically English noun-phrases: *fifteen years* and *seventeen years*. Note also how the use of English-language numerals is not restricted to a single participant: Todd, Lauren, Paul and Phoebe all use English-language numerals, with Phoebe even repeating Paul's use of "eighteen" directly, with rising intonation, as a way of seeking confirmation for his assertion. This repetition across a number of participants, then, can be taken to indicate that the use of these English numerals in this interaction is orderly and appropriate – within the current frame.

Considering the purist nature of Welsh standard language ideology, the use of English numbers in Welsh is not considered 'correct' language from a standard ideological perspective. But the practice of using English numerals in Welsh, especially for numbers above ten, is not unusual. Roberts (2000) attributes this practice to the Anglicisation of the Welsh education system – including arithmetic and mathematics – prior to the resurgence of Welsh-medium education in the second half of the 20th Century. For those who did not receive formal education through the medium of Welsh, it is assumed that counting is easier in English. But even amongst those who did receive formal education in Welsh, the use of English numerals is not unusual, if typically outside of 'standard' language domains. The use of English numerals in Extract 5, then, gives the interaction a more vernacular feel.

There are other examples of borrowings in this extract: *immaculate*, *hell*, *kitchen*, and *due*. Are these borrowings different in kind from borrowings found in the other extracts I have presented? What is significant, I feel, is that all of these words could have been substituted by other 'Welsh' words or phrases. I argue that these are not typically items that are used out of necessity, because of a lexical gap (although they may well be more easily accessible to participants), but

items where participants might reasonably be assumed to have a choice in their use. In some of these cases, it is clear that participants have a choice, since we find Welsh semantic alternatives within the same extract: *un-deg-saith mlynadd* ('seventeen years') in line 11, and *cegin* ('kitchen') in line 14, almost immediately after the same participant has used *kitchen*.

Can we assume that there is significance in the use of *kitchen* versus *cegin* here, and *seventeen years* versus *un-deg-saith mlynadd*? Within the context of this interaction, it is difficult to come to such a conclusion. In his use of *kitchen* and then *cegin* almost immediately afterwards, Todd doesn't seem to be doing any new or different social meaning work. We might infer that, since he uses *kitchen* in line 13 in quoting the tenant under discussion, he uses the more vernacular form, a borrowing, to enhance the performance. But I have interpreted that he is also quoting the tenant in his subsequent use of *cegin* (although it is less clear in this case that he is quoting). Moreover, there is no such possible interpretation in the contrast between his use of *seventeen years* (line 7) and *un-deg-saith mlynadd* (line 11). It would seem that in this interaction there is no significance to the individual use of a linguistic form.

However, what we might plausibly conclude is that within this interaction, participants do not seem to be making an effort to use language that is highly valued institutionally, perhaps as a consequence of their voicing frustration at the system and empathy with the tenant. They are not talking within an institutional frame. They are possibly even contesting institutional norms, albeit not consciously. As a consequence, the ideology of purism is not naturalised here, and even seems to be absent. I do not mean that this motivates the use of borrowings over 'Welsh' words. After all, we see the simultaneous use of *cegin* and *kitchen*, *un-deg-saith mlynadd* and *seventeen years*. Rather, borrowing is enabled. There is less value placed on keeping both languages apart (purism), but arguably more syncretic practice.

A similar example is found in *Voids*, where Pam is describing her experience with the local police, within the context of obtaining information about potential tenants.

“xx dwi’n cael fwy o **information** yn (place name) gynta nes i
gychwyn yn (place name) de (.) doth (.) yr heddlu (.) sy’n delio efo
(place name) yna (.) i (.) **introducio** nhw ei hunan i fi [...] a ma nhw’n
dod i fewn bob wsos (.) dwi’n cael mwy o (.) **wybodaeth** gynnon
nhw xxx”

(‘xx I get more information in (place name) as soon as I started at
place name) right (.) the police (.) who deal with (place name) there
came there (.) to (.) introduce themselves to me [...] and they come
in every week (.) I get more (.) information from them xxx’)

Like in Pam’s previous example (p. 226), we see that Pam’s narrative here, coloured with direct speech, is vivid. But Pam is also speaking with some passion, about a problem in their system that potentially has a negative impact on her own safety. This is another example of going outside of the main strategic institutional frame of a meeting to voice frustration. Again here we see the use of quite marked borrowings (*information* and *introducio*) for which there are fairly commonly used Welsh-language semantic equivalents (*gwybodaeth*, *cyflwyno*). Again, then, there is something of a vernacular feel to this example. Potentially, this vernacularity is used in voicing frustration – as well as, in this example, in painting a vivid picture of a previous conversation. But again, we see a Welsh-language semantic equivalent for one of the borrowings, *gwybodaeth* (‘information’), in the very same utterance. This suggests that it is not necessarily that borrowings are given precedence by Pam over ‘Welsh’ words, but that the need to keep both languages separated is suspended: the absence of the ideology of purism.

Conclusions

What I suggest, then, is that borrowings and not-codified terms are seen as appropriate to some extent. This would suggest that the ideology of purism – which underpins lexical planning – isn't entirely naturalised at the Association, and that to some extent, it is in competition with other interactional priorities. These include, for example, whether lexical items are primed in the local context, how easily and quickly the speaker can access lexical items in talk, and the speaker's judgements of recipient familiarity (these last two points are related to priming, but also to speaker and recipient competencies). However, some qualifications need to be made. It is not clear that *all* borrowings and not-codified terms would be considered appropriate. If purism isn't naturalised, that does not mean to say that all borrowings would be considered normal or appropriate in talk that is made mostly of Welsh-language resources, after all. I suggest that there is greater tolerance for borrowing in the case of words that are difficult to translate into Welsh (such as *manual handling*, *utilities*, *issue*), or words that are less salient within the organisation. If lexical planning targets primarily 'new' words (as we might interpret Terminology to do, if not lexicography), then it would seem that lexical planners' priorities do not overlap entirely with participants'.

However, there seems to be some variability in the extent that borrowing is seen as appropriate and in the prioritisation of purism. We find that purism is most highly valued in the *Minafon* recording, which I attribute to the pressure on participants in this meeting to perform through the medium of a code that is highly valued at the institutional level. In Chapter 4, I asked, given that the Association organises itself as a hierarchy, whether there was a kind of Welsh that is highly valued at the institutional level and, if so, whether it overlaps with the symbolic code that lexical planners seek to spread. My interpretation of the *Minafon* recording would suggest that 'pure' Welsh – born of an ideology that, I argue, does overlap with lexical planners' – is indeed more highly valued institutionally. And yet, the same proviso applies here: that this is only true to the

extent that it does not interfere with the business of the Association. These meetings are functional, of course, *Minafon* in particular, with participants quite determined in their pursuit of the business priorities of the Association.

By contrast, we found in the case of the other meetings, *Estates* in particular but to some extent in *Voids*, that purism was not only in competition with other interactional priorities, but was potentially suspended. In Pam's case, I suggested that purism would not allow her to colour her talk in the way she does. This would suggest that in *Voids*, purism was quite a low priority. Similarly, in Nell's use of *taleb* and *voucher* (also in *Voids*), other than in direct reference to written documentation, it is the more vernacular and accessible *voucher* that she uses when she most needs to make herself understood. In the case of the frustration that is keyed in the *Estates* meeting, it is possible that institutionally-valued purism is suspended as a means of voicing that frustration. I have emphasised that this particular interpretation is rather uncertain. And yet it is fruitful to think back to the sociolinguistic context that I outlined in Chapter 4, where I suggested that the use of borrowings was more typical of vernacular ways of speaking, and that 'standard' (which would include 'pure') Welsh can be read with a class/status group meaning. What we might be seeing in the *Estates* recording, then, is a case of vernacular maintenance (Milroy & Milroy, 1997: 53). On one level within the organisation, then, we may be seeing a rejection of purist ideology.

CHAPTER SEVEN:

Conclusions and implications

7.1 WHETHER, HOW AND WHY? ADDRESSING MY RESEARCH QUESTIONS

Earlier in this thesis, I laid out some research questions to ask in relation to the lexical planning endeavour in Wales (p.45). Within the settings that I had selected, I set out to ask whether Welsh speakers use planned lexis, what social meanings are activated by the use of planned lexis, and what constrains the use of planned lexis in interaction. In this chapter, I consider these questions in light of the analysis of both datasets together. I also evaluate the methodology I used in seeking to answer these questions. Finally, I consider my analyses and findings against the wider sociocultural framework of language standardisation, language planning and revitalisation that I discussed at the beginning of the thesis.

One of the difficulties that my research questions have come up against is how to define planned lexis. To some extent, my two different analytic approaches tackle this problem in different ways. My first, a quantitative methodology, sought to identify subject-specific or work-related vocabulary in the data and to compare that against lexical codification texts. However, it was problematic even to define what was subject-specific or work-related. As I asked in Chapter 6, could we reasonably expect that the use of *tir* ('land') by participants was of any interest to lexical planners? My other approach sought not to focus specifically on subject-specific or work-related vocabulary, but to consider all vocabulary that could be considered lexical items (carrying significant semantic content) through the lens of purism and standard language ideology, and to ask whether these words – borrowings or not – were presented as normal in interaction. This approach, then, was a more ideology-oriented approach than the more technical first approach.

Taking both analysis chapters together, the quantitative analyses of codification

suggested that participants did use planned lexis, even a majority of the time. Yet use of planned lexis is qualified, since a fair amount of the terms identified were categorised as not codified. Despite this apparent similarity between datasets, it is worth noting that there were differences: the proportion of codified terms in the *Syrjeri* data (45%) was substantially lower than the proportion of codified terms in the Housing Association data (60%), suggesting perhaps that participants in the *Syrjeri* data follow less clearly the stipulations of lexical planners. Somewhat contrary to this conclusion, however, the proportion of not-codified terms in the *Syrjeri* data (16%) was a fair amount lower than the same category in the Housing Association data (25%). The difference here is caused by the composition of the more ambiguous categories: the *Syrjeri* data had a higher proportion of terms in the partly codified, mixed and not applicable categories than the Housing Association data. It is not necessarily the case, then, that talk in the Housing Association is more clearly 'planned', but that there is more ambiguity in the lexical planning endeavour in the field of health. Perhaps this is surprising, considering that much more work has been carried out in lexical planning in this field than in the fields related to the Housing Association's work. And yet, it is also possibly the case that it is simply a more difficult field to plan for.

It is also worth mentioning that the terms in the *Syrjeri* data that I interpreted to be more specialist were typically found in the not-codified and not-applicable categories. (The terms in the Housing Association dataset did not seem particularly specialist at all.) If lexical planners are truly interested in specialist terms, then this finding would suggest that participants did not follow lexical codification texts. However, I have suggested earlier that I am sceptical about the specialist focus in lexical planning – that planners target everyday language practices. The problem here, of course, is knowing exactly what planners are interested in, how to define it and how to analyse it.

When we look at the issue through the lens of purism and standard language

ideology, however, we get a slightly different view. Purism seems to be a fairly prevalent ideology in both datasets. We see from the policy analysis for the *Syrjeri* data, for example, that purism is presented as highly valued within the institution, with appeals to ‘correct’ language and ‘rich’ vocabulary, and to restrict the use of literal translations of English idioms. This ideology seems to be borne out to some extent in the data itself, with participants – presenters and Dr Ann, at least – using ‘Welsh’ words even where they seem to judge lexical content is potentially inaccessible and requiring an English gloss. In the Housing Association data too, purist ideology seems to have a certain influence, as evidenced by the use of markedly ‘Welsh’ lexical resources for quite salient objects of knowledge within the institution, as well as some evidence of self-repair, replacing more ‘English’ forms with more ‘Welsh’ ones (as we saw in the case of *checio* and *profi*).

When we consider the question of whether planned lexis is used in each dataset, then, we can conclude that yes, to some extent, it is – either gauged by comparing lexical items used with the stipulations of lexical codification texts, or viewed through the lens of the prevalence of purism as an ideology.

However, in both datasets we saw linguistic variation, both contact-induced variation and variation in the distribution of terms according to codification category. Contact-induced lexical variation refers to the presence in a single dataset of semantically equivalent lexical items that might be interpreted as ‘Welsh’ and ‘English’ (although in each case, the lexical item might, etymologically speaking, originate from another language). These language-differentiated lexical items were typically found in different codification categories, with ‘Welsh’ terms either found in the codified or mixed (accepted by some texts, not by others) categories and ‘English’ in the not-codified category. This variation in each dataset suggests that the implementation of lexical planning is not merely a matter of speakers using a planned word or not, since it points to differentiated practice.

The question of differentiated practice relates to my research question regarding the social meanings activated by the use of planned lexis, or not, in the spoken data. Further analysis of the quantitative data, breaking the data down according to participant type in the case of the *Syrjери* data and meeting in the Housing Association data, sheds some light on differentiated practice (although not specifically on the point of lexical variation). It was suggested that lexical planning might be differently implemented according to participant in the *Syrjери* data, and according to meeting in the Housing Association data. Both these results point in some way to the roles of participants within the interaction. In the *Syrjери* data there were clearly differentiated roles in the interaction, and the use of planned language seemed to vary with these roles. More indirectly, perhaps, in the Housing Association data, differentiated practice according to meeting might be interpreted as resulting from the different roles played within the interaction. However, it is not clear whether differences according to role might be caused by the topics dealt with by participants (topic necessarily has an influence on lexical content), or whether differences are caused by differentiated norms and assumptions of appropriateness.

Nonetheless, I have questioned what exactly distributional patterns tell us about the social meaning activated by the use of lexis in interaction – whether we can read off social meaning from comparing frequencies of linguistic forms according to independent variable, and whether quantitative data can adequately account for all the meaningful variability of talk (see Chapter 3). In the qualitative analyses I concluded that, on the whole, the lexical items that were used were typically presented as normal within institutional discourse. There was very little playing with words in the *Syrjери* data, with the notable exception of the word *syrjери* itself. Nonetheless, we must question what the apparent avoidance of Welsh words in some cases, as well as the presentation of Welsh words side-by-side with borrowings, tells us about social meaning. I have argued that participants – presenters and Dr Ann in particular – conceivably avoid certain

'Welsh' lexical items, or reformulate them, in order to ensure that their talk is accessible (a requirement made in S4C's language guidelines). But what I argue is that orienting to accessibility is motivated by a desire/perceived need to communicate on a fairly equal footing with the speaker. In orienting to accessibility, the speaker takes the communicative burden (Lippi-Green, 1997) on themselves, rather than placing it on the recipient. The social meaning doesn't necessarily directly pertain to the linguistic form itself, but to the identities the participants perform in interaction and to their relational goals. Arguably, if participants use language that is too inaccessible, this might put too much social distance between themselves and their recipients. I argue this is particularly true where participants have previously had time to reflect on language practices (as is the case of Dr Ann and the presenters), since it eliminates, to a large extent, the question of the speaker's own linguistic repertoire.

One of the most significant interpretations from the Housing Association analysis in terms of social meaning was that 'pure' Welsh did indeed seem to be valued within the institution, suggested by the greater naturalisation of purism within the high-stakes *Minafon* recording. However, 'pure' Welsh was not completely naturalised within institutional talk. Confronted with the need to pursue institutional goals, purism faced a challenge when it came to lower-frequency objects of knowledge, which are probably more easily accessed by means of their borrowed designations. In contrast with the partial valuing of 'pure' Welsh, the relaxing or suspension of purism seemed to be used as a means of challenging the institutional narrative. The borrowing of fairly common words – from quite basic vocabulary – gave interaction a vernacular feel, and by so doing contrasted with the norms of the institution to voice empathy and frustration.

This discussion of social meaning is closely connected to the question of constraints on the implementation of lexical planning. Arguably, if participants avoid the use of planned language because of its social meaning implications, this places a constraint on the implementation of lexical planning. Is this of concern to

lexical planners? This depends on the kind of genres and registers that lexical planners are targeting. Arguably, in the case of the more vernacular-style practices in the Housing Association data, lexical planners might not consider this kind of talk to be within their scope. On the other hand, what my analyses do emphasise is that sometimes participants do seem to prioritise other considerations, such as the accessibility of talk, and this can occur across the spectrum of genres and registers. Perhaps it is best to consider that there are a number of local contingencies that affect language practice in situ, and that language planning is one of these, but by no means exclusive.

I have also noted some other constraints to the implementation of lexical planning. In my quantitative analysis of codification, what the more ambiguous categories in both datasets point to is that it is difficult to stipulate entirely how language should be used. Codification is not amenable to close definition. There are a number of reasons for this, including disagreement between lexical planners, the indeterminacy of register boundaries, and the infinite nature of language (you cannot plan all of language, because there is no limit on what can be done with language). It might be argued that this indeterminacy might make it difficult for the users of lexical codification texts to know how to interpret those texts.

7.2 METHODOLOGICAL MATTERS

It is worth pausing here to consider and evaluate the methodology used and hence the epistemological grounds upon which I seek to make my claims.

In terms of the overall focus of the thesis, although it is nominally concerned with lexical planning, and what some planners would call ‘standardising terms’ or ‘terminology’, I have also focused quite significantly on borrowing and purism. This opened up the scope, not only to look at what might be called ‘terminology’, but to look at all kinds of lexical resources. Is it possible that the study has less to say about ‘terminology’, then, than planners might have expected? I followed this

route for a number of reasons, not least because of the difficulty of determining the focus of the lexical planning endeavour. Moreover, as I stressed in the first chapters of this thesis, the lexical planning endeavour sits quite firmly within a discourse of purism, and is as much concerned with controlling the use of borrowings as with coining and standardising new ‘Welsh’ words. Controlling borrowings and coining new ‘Welsh’ words are two sides of the same coin. I consequently argue that it is difficult to see a difference in kind between using a borrowing for a ‘new’ concept and using a borrowing for an ‘old’ or well-established one. The focus on borrowings was borne out to some extent in the quantitative analyses. It seemed quite clear from these analyses that, typically, the not-codified category consisted of borrowings, especially marked borrowings. There wasn’t a clear mapping of borrowings onto the not-codified category, but borrowing, especially without integration, emerged as a key theme. I am not arguing that there are not other issues that relate to the lexical planning effort besides borrowing and matters relating to bilingualism. For example, there is the question of whether there is disagreement amongst lexical planners or not, and if so, whether speakers show preference for one over the other. But the main issue that emerged from the data itself was purism. We saw this, for example, in the *Syrjeri* data, where participants showed preference for using ‘Welsh’ forms even though they were not used in a manner that was consistent with lexical codification texts, or even where the forms were not found in lexical codification texts.

To focus on some methodological matters concerning the specifics of the research choices I made, it is possible that I would have gained further clarity on the Housing Association data had I considered looking at the distributional pattern of codification categories according to participant or participant type. One possibility would have been to group participants according to their level in the organisational hierarchy, and so to consider the distributional pattern of terms according to whether staff members were higher-ranking or lower-ranking staff. This was a category that I typically felt was salient in the data in any case

(and that I tended to treat as salient in the qualitative analysis). I decided against taking this route, however, not least because there were so many participants in the data. One difficulty with the sheer number of participants was determining the identity of the speaker in transcribing. I often found that establishing the identity behind the voice was a matter of guesswork, which made it difficult to categorise data according to participant. I also considered comparing the lexical use of a smaller number of participants. Some participants have quite distinctive voices and ways of speaking, making them easier to identify. However, this would have meant limiting the amount of data open to this part of the analysis, leaving a data sample that I felt was too small.

In terms of my qualitative analysis, there is inevitably a degree of indeterminacy. This is a common problem in qualitative research, that “there is a substantial risk that we simply use intuition, casual observation, or theoretically motivated observation to draw conclusions, thus importing conceptual constructions that cannot be substantiated in the data” (Mäkitalo & Säljö, 2002: 60). Stroud (1992) has made similar observations in relation to understanding the social meaning of code-switching. He critiques code-switching research that, in his view, made too definitive interpretations of participants’ motives for code-switching. He argues that researchers have no way of knowing whether the researcher’s interpretation would be valid to the participants themselves. That is, researchers tend to be too hasty in inferring speaker intent. For Woolard (2005: no p.n.)

“Stroud is right to call into question our preference for tidy accounts over ambiguity, to ask that we provide warrants for interpretations, and to highlight the perennial problem of falsifiability. But these problems are only as insurmountable for codeswitching as for all other topics that concern the interpersonal import of human activity and the conclusions that people draw about the actions of those around them, that is, the stuff of anthropology.”

With this in mind, I have sought to inform my analysis by referring to other sources, such as policy documents and the sociolinguistic literature on Wales. Where these sources have ceded little insight, I have sought evidence from the data itself, or I have sought alternative interpretations and inferred the most plausible. But there inevitably remains a degree of indeterminacy.

Nonetheless, it might be argued that indeterminacy in qualitative analysis is, at least, more transparent than indeterminacy in quantitative analysis. In the latter, analytic categories can seem like fixed objects – indeed, need to be treated as such – whereas they are, in the case of language, human constructs. They are ideas more than they are things, and as such they are malleable and subjective. It is evident in my quantitative analysis that I have had to pigeon-hole a lot of data into categories that are very difficult to define. In some cases it was easy to determine whether a word was ‘codified’ or not, according to my criteria. But in others the waters became cloudy, and terms found their way into the mixed or partly codified categories. But even in the case of the codified category, it is a construct that I defined by selecting a handful of lexical codification texts out of a number of possible texts, and without knowing exactly what texts were known to participants. Moreover, there was inevitably indeterminacy in my selection of lexical items as terms for analysis. It was difficult to determine what was health- or work-related, and even what was a word. The indeterminacy in my qualitative analysis, then, may be no worse than the indeterminacy in my quantitative analysis.

There is an epistemological tension at the heart of all sociolinguistic research. This tension is likely inevitable, considering the focus of sociolinguistic research is so subjective and continuously shifting. The answer, or the best approach to research, is perhaps not to focus on one research method, nor even to search for and accept definitive answers, but to use multiple research methods in order to gain different perspectives (Coupland, 2007). What I have offered, then, is an interpretation from two different perspectives. I do not claim to have presented a

definitive analysis, but invite the reader to consider my interpretation critically.

7.3 SO WHAT? IMPLICATIONS FOR LANGUAGE PLANNING

At the beginning of this thesis, I presented the lexical planning effort in Wales in the wider sociocultural framework of language planning, arguing that current efforts at lexical planning were envisaged as part of the revitalisation effort for Welsh. In this section, I consider the wider implications of my study for the lexical planning effort as well as for revitalisation more broadly.

Following from the assumption that the lack of use of Welsh in certain domains has hampered the development of the Welsh lexicon, which further discourages the use of Welsh, lexical planning is ostensibly carried out as a remedy for this failing. Primarily, it is taken that “the extension of the register range of Welsh means mainly a series of decisions concerning a specialized terminology to be recommended” (Thomas, 1982: 91), quite a monologic conception¹⁵⁷. It is perhaps not surprising that this study highlights that this assumption is oversimplified, that you cannot simply make a series of decisions regarding linguistic forms and meanings and expect speakers to use that language unquestioned. In fact, I suggested as much in my preliminary chapters, and my analysis confirms that there are number of considerations that might influence the take-up of planned language. These include differences amongst lexical codification texts and potential difficulties interpreting them, access to planned language (its dissemination, speakers’/recipients’ repertoires and in situ scope for reflexivity), as well as more sociolinguistic concerns such as identity, genre and ideology.

Nonetheless, that does not mean to say that the lexical planning effort is considered futile or worthless. Although it was difficult to judge whether

¹⁵⁷ It is fair to emphasise that more contemporary lexical planning – term planning specifically – seeks to work with various groups in order to improve the relevance and take-up of lexical planning projects. That is, the process is somewhat more dialogic. However, the product can still be described as an authoritative statement of what is and is not considered correct by planners.

participants used the planned lexis itself (partly because of the difficulty of defining planned lexis), I would argue that lexical planning isn't completely out of tune with much of the practices observed in the data. That is, the ideology of purism, being quite central to lexical planning, was also quite prevalent in the data, if not completely naturalised. This was particularly the case in the *Syrjeri* data, but also to some extent at the Housing Association. That planning is subject to local contingencies is not to say that planners are completely out of touch. It does not seem to be the case that lexical planners are seeking to impose a language from a position wholly alien to other Welsh speakers, something that Fishman (1983) warns against when he notes that corpus planners can become "butts of humor, sarcasm, and ridicule, unappreciated at best and vilified at worst." (1983: 116). It is hardly surprising that lexical planning is, to a certain extent, in touch with practices observed in the data, given the long history of lexical planning in Wales. Purism and standard language ideology have long been culturally transmitted, and they are not only imposed from above – 'above', in this case, referring to lexical planners.

Nonetheless, I would argue that these ideologies are not culturally transmitted wholesale. To some extent, it is at the institutional level that they are most prevalent, and there is reason to believe that, in more vernacular practice, purism and standard language ideology are much less prevalent, that lexical planning does not affect to the same extent the everyday language practices of Welsh speakers outside of institutional settings. Jilg's (2003) finding that differentiated social networks have differentiated lexical resources suggests as much (see Chapter 4). Whilst it is difficult to infer that there is a clear class difference between Jilg's social networks (minority language groups are not typically easily defined in relation to class (Williams, 1987)), they might be described as representative of different status groups. In my own data, there is a suggestion that, where participants play out identities that might be defined as divergent from the institutional narrative, purism and standard language ideologies are suspended, possibly even contested (although not consciously). In

this light, it is important to emphasise that it is primarily the institutional level that I have observed, and it is not clear from my analysis to what extent lexical planning is in tune with other types of discourse. We should not be too hasty to conclude that lexical planning is in tune with the ideologies of *all* Welsh speakers. Nonetheless, as I argued in Chapter 2, it is at the institutional level that standardisation will be achieved, if at all, and so it is not insignificant that purism and standard language ideology are fairly prevalent within the institutions observed.

Whether or not lexical planning is broadly in keeping with on-the-ground ideologies, there remains the question of whether lexical planning ought to be reimagined, considering the local contingencies that hinder its full implementation. Whilst some of these can be viewed as management issues (how well lexical codification texts are disseminated, centralisation of planning, etc.), the sociolinguistic issues of genre, identity, ideology, etc. are more difficult to plan around. Moreover, they raise question regarding the aims of lexical planning in the first place.

Reimagining can be, and is, discussed not only in the context of lexical planning, but of corpus planning and standardisation more broadly. I outlined in Chapter 4 that there has been a debate regarding the nature of ‘standard’ Welsh, which is often characterised as deriving from the fact that ‘standard’ Welsh has been esoteric and perceived to be too far removed from everyday practices. The debate still continues, as is evidenced in two recent articles published in the online magazine of a Wales-focused think-tank (the Institute of Welsh Affairs’ *Click on Wales*). David (2013) called for the reform and simplification of Welsh, under the leadership of a Welsh language academy, on the basis that Welsh is too ‘difficult’ that it is not ‘fit’ for modern purposes:

“Welsh is a hard language to learn, as legions of people throughout the past century who have learnt it in school but who emerged

barely able to put a sentence together, and the equally large number of drop-outs from adult education and home learning also testify.”
(no p.n.)

Even first-language speakers, he claims, lack confidence, and he cites the ubiquitous “My Welsh isn’t good enough” as “the commonest refrain in Wales”. Whilst my research has focused primarily on lexis and purism, I consider these matters to be part of the broader issue of linguistic repertoires and ideology, and hence closely related to what David refers to as the problem of ‘difficulty’. Moreover, as I argued in Chapter 5, the perceived need to express less common concepts in ‘Welsh’ can impose a cognitive burden on interactants, because of the weakening, through translation, of the intertextual link between the form used in a specific interaction and the historic habitual practice that it refers back to through partial repetition (a translation being a repetition of semantic meaning but not of form). In David’s words, it is more ‘difficult’ to decipher semantic meaning where the words used are not common¹⁵⁸.

In his response article, D. G. Jones (2013) dismisses David’s arguments as ideological. In relation to some of David’s linguistic points, I tend to agree with Jones. David compares ‘difficult’ Welsh unfavourably with English, “a stripped down language that is constantly simplifying itself”. Is Welsh ‘difficult’ or is it merely different from English? And what does it mean to say that English is ‘stripped down’ in any case? Moreover, in arguing that contractions in spoken usage make Welsh difficult to understand, David claims “It is as if Cilla Black had

¹⁵⁸ It might well be argued that this ‘difficulty’ is a consequence of the social, economic and political conditions under which Welsh is practiced – that ‘Welsh’ words would gain currency if Welsh was practiced more widely, something that is restricted under the current arrangements. Indeed, some new Welsh words do seem to have gained currency, making their deployment in practice less difficult. But I remind the reader that ‘difficulty’ has, in the past, been highly valued within Welsh standard language ideology, where ‘standard’ Welsh was envisaged as a special skill to be mastered (see Chapter 4). It was a means of showcasing one’s mental capacities (education) and status (see Coupland & Kristiansen’s (2011) discussion, drawing on Bourdieu, on *distinction* in relation to standard language ideology). Ideology and practice has arguably changed significantly since the early 20th century, but it is certainly a possibility that remnants of this aspect of standard language ideology remain – that lexical planning is planning in difficulty on ideological grounds.

been put in charge of English and we had all been taught “worra lorra” or “you’ve gorra lorra” were the new correct forms of speech”¹⁵⁹. In the context of a sociolinguistic thesis, I will assume that the contentiousness of this statement does not need explaining.

Nonetheless, underneath the confusion of David’s arguments, I do believe there lies a problem that needs to be addressed. I feel that both David and Jones are missing the point, David because he misinterprets it, and Jones because he doesn’t see it (possibly because of his own ideology). “My Welsh isn’t good enough” is a common refrain. David describes the problem as Welsh being too difficult. I would describe the problem as a mismatch between purism and standard language ideology on the one hand, and the linguistic repertoires of a substantial proportion of Welsh speakers on the other. Linguistic repertoires cannot live up to these ideologies for many reasons, including language shift and the restricted practice of Welsh in various domains. Incidentally, I would argue that these ideologies can confound the problem of restricted practice, since they can discourage those who feel limited in their ‘Welsh’ linguistic repertoire from engaging in various domains through the medium of Welsh. Whilst purism is not all-encompassing (there are plenty of borrowings that are felt to be ‘standard’ and ‘legitimate’) arguably, the extent of purism as an ideology places further pressures on a linguistic culture that is already struggling. This is, in fact, a reframing of the ‘lexical gap’ argument. Whilst it is commonly assumed that Welsh speakers can’t engage in certain domains because they lack the vocabulary, I argue that an ideologically-motivated sense of appropriateness potentially deters engagement in certain domains without appropriate linguistic resources, notably vocabulary. From this perspective, the solution doesn’t have to be forging new linguistic resources, but shifting people’s thinking.

Taken to its logical conclusion, purism requires that all ideas expressible in

¹⁵⁹ Cilla Black is a celebrity from Liverpool. *Worra lorra laughs* (‘what a lot of laughs’) is a famous refrain of hers. The Liverpoolian accent is one of the most highly stigmatised of British accents (see, for example, Bishop, Coupland, & Garrett, 2005).

English can also be expressed using Welsh linguistic resources. Lexical planners have been at pains – for centuries – to demonstrate that Welsh is copious, “that anything which can be said in English can, in so far as a dictionary can demonstrate, be said in Welsh. Or to put it differently, that Welsh is a language like any other” (Jones, 2013). Lexical planning, then, is seen as a means of legitimising Welsh, by making it fit into the dominant structuralist ideology of language, for a language that has to borrow is not a complete structure.

But the measure used to prove the copiousness of Welsh has been to coin and codify ‘Welsh’ semantic equivalents – often translations – of English words. Welsh might fit into the structuralist idea of what a language is as a consequence of lexical planning, from having it proven that ‘native’ Welsh resources can be combined to express any concept imaginable. But is this feasibly evidence of copiousness? Or is it merely a show of productive potential – and a somewhat defensive one? Arguably, if speakers don’t use this codified language, it does not serve as evidence of copiousness. Is it not from practice that we should judge what a language is, and not from a catalogue of decontextualised linguistic forms? Mugglestone (2011: 77) takes this view when she argues that real linguistic authority rests in language practices, and not in the pronouncements of prescriptive lexicographers:

“the real processes of legitimization and citizenship, whether of words or meaning, instead take place in usage (and evidence of this). Power here comes from the masses. It is authority of this kind, based in the pull of language practice by the collective of users – the reality, in other words, of what people actually do in using a given language – which serves to explain the dominance of some loanwords (*le weekend, l’é-mail, le spam*) against the loss of others. Even if official documents can perhaps be made to toe the official line, policing everyday acts of usage is much more difficult.”

I do not agree entirely with this view. Centralised, institutional authority can have a powerful effect on language ideologies and language practices. However, where there is conflict between prescription and practice, practice is primary. In the early days of Welsh lexical planning, William Owen Pughe coined numerous words, some of which had no practical use. These included, for example, *draeneta* ('to hunt hedgehogs') and *damzistryciad* ('a being foaming, or spraying all round') (cited in Morgan, 2002). Moreover, these words were constructed following a method that was grounded in an ill-informed etymology, as well as represented through a very unconventional orthography (see Chapter 4). Both form and function were defunct, and these words did not catch on. Despite being codified, catalogued as 'Welsh' in Pughe's dictionary, they go no way to establishing the copiousness of Welsh, because they are not practiced.

This discussion, of course, relates to the very nature of language. Language planning and policy is typically embedded within a structuralist perspective, the dominant perspective on language in the Western tradition, that sees language as a structure that foreshadows practice, and practice as the deployment of this structure, and independent of it. An alternative view is that structure emerges from the repeated practice of language (Pennycook, 2010). From this perspective, there is no *one* language, but many, although there emerge shared ideas about what language is, and what counts as belonging to a language¹⁶⁰. That is, language emerges as a cultural construct "consistent with cultural beliefs, norms, and values that are presumed to be shared by others in the local situation and perhaps more broadly by actors in a broader community (e.g., the organization or society)" (Johnson, Dowd, & Ridgeway, 2006: 57). The cultural construct of language might be well-established (in any given culture), but from a language as practice perspective, it becomes difficult to justify the view of a language as a discrete bounded structure that has any objective reality.

¹⁶⁰ Although it is by no means clear that a socially shared conception of a common language is inevitable (see, for example, Grace's (1990) work on Melanesian languages).

Wright (2007) points out that the language-as-practice view has hardly surfaced in language policy, that policymakers typically disregard linguists' claims that it is impossible to define languages unequivocally as determinate, bounded objects. Wright sees this as inevitable to some extent, arguing that a language policy that treats language practices as constantly shifting would be too difficult to manage across institutions. Even if one were to seek to implement a policy grounded, in theory, in the recognition of language as practice,

“In practice, there may be little difference because of the difficulties of translating recognition of diversity into policy. So, even if we conceive language as contextually bound performance and ideologically grounded practice, we tend to implement policy as if language were an ideal system. The fundamental dilemma of the practical implementation of linguistic rights lies here: how to reconcile political exigency, which tends to be centripetal, with actual language practice, which tends to be centrifugal?” (Wright, 2007: 212)

Wright may be right that there are difficulties in applying a view of language as practice in a policy framework. Nonetheless, there are approaches in sociolinguistics that seek to challenge the structuralist premise at the foundation of traditional LPP. For example, in Corsica “academics promoted the idea of Corsican as a “polynomic” language defined both by its internal variation (multiple centres of “authenticity” and “authority”) and by speakers’ recognition of linguistic unity in diversity – a collective stance vis-à-vis linguistic variation that challenges the very principles of dominant (French) language ideologies in its inclusive, non-hierarchical nature” (Jaffe, 2003: 516). Insisting on language as polynomic, then, would open the scope to tolerating linguistic variation – that is, shifting thinking rather than seeking to directly influence linguistic form (corpus planning as we know it). In Corsica this shift was pursued through education policies. Jaffe notes that polynomy doesn't embrace contact-induced language

variation, but I wonder whether she is categorically stating that it cannot, or merely observing that it has not done so in the case of Corsican. She is also somewhat sceptical regarding the viability of putting a pluralistic model of the minority language (Corsican) into practice, because of the contrast with the way French, the majority language, is taught:

“a historically-embedded practice that took a single, authoritative norm absolutely for granted. [...] maintaining a polynomic perspective on standards in Corsican while enforcing a single French standard runs the risk of conveying the message that Corsican is less authoritative – less a legitimate linguistic system – than French. This is a message that most bilingual Corsican teachers would be loathe [sic] to send, since they view their schools as showcases for Corsican language status” (Jaffe, 2003: 536)

Nonetheless, polynomy does highlight how policy might look if it accepted the language-as-practice perspective.

Another example of where a view of language-as-practice might be implemented as policy is Fairclough’s (1995a) critical language awareness. In the context of English in the United Kingdom, Fairclough advocates teaching children (and others) to think critically about language – to understand something of the historical processes that have led to the powerful position of standard English – and to value their own language varieties, not as some long-dead museum artefacts to be preserved for the sake of ‘heritage’, but as living media of expression and communication/community. Fairclough doesn’t argue against the teaching of ‘standard’ English, since standard language ideology in the United Kingdom is so prevalent that to *not* teach pupils ‘standard’ English would be to do them a disservice. But teaching critical language awareness would seek to ensure that pupils emerged from formal education with the ability to compete in

a modern economy, without the standard ideology being so naturalised as to be damaging to vernacular speakers of varieties quite different from the ‘standard’.

In how far does my own research suggest that such approaches might be desirable? Does lexical planning *need* to be reimagined? I emphasise that it isn’t entirely clear in the data that the lexical planning effort is problematic. Whilst at times lexical planning isn’t followed, or purism isn’t wholly naturalised in the interaction (sometimes not at all, and potentially contested), there is considerable overlap between the ideologies of lexical planners and those found in the data. Moreover, it isn’t clear that purist or standard language ideology deters people from contributing. This is something that is difficult to detect with the methodology I have used. Interaction analysis largely focuses on who is present, not who is absent or why. But the analysis does indicate that where the implementation of lexical planning is potentially problematic – for example, where the use of a ‘Welsh’ word might make the referent less easily accessible – participants typically implement their own covert language policies. Moreover, S4C’s formal language policy explicitly makes room for those whose linguistic repertoires might not overlap with what S4C would describe as ‘correct’ or ‘rich’ Welsh. There is, then, a degree of tolerance of linguistic variation which, I would argue, is also seen in the Housing Association, where even those at the very top use at least some borrowings, sometimes marked ones. What we might say, then, is that at the institutional level, lexical planning is implemented where possible, but that room is made for other ways of speaking.

It might be argued, moreover, that such a situation is inevitable. In Chapter 2, we saw that the French state, for example, has pursued standardisation and purism quite fiercely, making the use of borrowings punishable as fraud. And yet Weinstein (1989) and Cartrite (2009) were sceptical as to the efficacy of these measures, stringent as they might seem. I do not expect to see similar measures implemented in Wales in the foreseeable future, partly because lexical planning seems to be of low priority politically, relative to other language planning

matters. However, the French example serves as a reminder that even quite heavy-handed lexical planning and enforcement measures do not necessarily lead to outright public acceptance. If this is how lexical planning is implemented on the ground, is there a need for reimagining?

In some ways, this is not a question that can be answered through recourse to interaction data, but a philosophical question. In terms of purism, for example, Shapiro (1989: 28) views purism quite negatively:

“A purification movement is representative of a centripetal tendency. Diversity is under attack, and the discursive economies that result can only enhance a centralizing tendency within the society’s system of power and authority.”

But I feel purism can also be seen as a response to a perceived attack on diversity. Purists can see themselves as opposing social movements that are political in themselves, as protecting diversity by insisting on cultural (re)production in languages or through linguistic resources that have been minoritised as a consequence of social, economic and political conditions. Purism may well be a centripetal tendency. But it is not necessarily in opposition to a centrifugal tendency, but to another centripetal one (in the case of minority language struggle), one that is construed as inauthentic. It could be argued that this has been the story of Welsh purism, from the 16th century to the present.

Viewed in this light, Welsh lexical planning might be viewed as a positive and empowering step in the revitalisation effort. However, whilst I have emphasised that there is a degree of tolerance of variable language practices in the data, there is a suggestion that certain institutional roles are closed to those who do not speak a certain way. This is quite clear in the media study, with S4C’s language guidelines specifically requiring that presenters speak ‘correctly’, and the spoken data apparently mirroring this policy. That is, although S4C welcomes all Welsh

speakers into its programming, regardless of how they speak, it only welcomes everybody for marginal roles. In the Housing Association, it was less clear that access to institutional roles was limited to those who speak ‘correct’ or ‘standard’ Welsh (or otherwise defined). But there was a contrast between some of those who inhabited the higher-ranking posts and some of those lower in the hierarchy, as well as more clear signs of metalinguistic awareness of ‘correctness’ in the *Minafon* meeting, which involved more high-ranking staff and where stakes seemed to be higher.

Wright (2007) has highlighted that minority language rights movements face a paradox in standardisation. Minority language groups seek empowerment through language rights. However, the resulting need for standardisation (Wright considers that standardisation is an inevitable consequence of the introduction of positive language rights¹⁶¹) potentially results in placing pressures on those who do not speak the ‘standard’ to converge:

“Where language rights are granted, it is likely they will be the right to employ an ideal, agreed system rather than the right to communicate with the institutions of the state in one’s language of primary socialisation. This leads to the situation where there is pressure on those speakers who do not use the standard employed to converge, to shift to that norm.” (Wright, 2007: 211-2)

Moreover minority language rights movements, in seeking legitimacy, can inadvertently create their own minorities, thus reproducing “the centre-periphery conflict that accompanied the spread of national languages” (Wright, 2007: 212). The role differentiation that I have argued is found in the data potentially points to such minoritisation in the case of Welsh.

¹⁶¹ ‘Positive rights’ refers to the use of a minority language in administration, in justice, as a medium of education, etc. where the state is required to take action to ensure that services are provided in the minority language.

This situation is paradoxical and, from a critical perspective, problematic. Even from a more practical standpoint, double minoritisation might be counterproductive to the revitalisation effort. If a minority language speaker is also a fluent speaker of the majority language – as is the case for practically all Welsh speakers – they have a choice not to align with the minority at all, but to shift to the majority language. Whilst lexical planners have argued that a *lack* of ‘Welsh’ words has driven people to ‘abandon’ their mother tongue (see Chapter 3, section 3.5), and have therefore been spurred on to further the lexical planning effort, it is possible that too much standardisation is a confounding factor in language shift. This may well be a reason to reimagine lexical planning.

References

- Agha, A. (2000). Register. *Journal of Linguistic Anthropology*, 9(1-2), 216-219.
- Aitchison, J., & Carter, H. (2004). *Spreading the Word: The Welsh language 2001*. Talybont: Y Lolfa.
- Alfonzetti, G. (1998). The conversational dimension in code-switching between Italian and dialect in Sicily. In P. Auer (Ed.), *Code-switching in Conversation: Language, interaction and identity* (pp. 180-214). London: Routledge.
- Alvarez-Cáccamo, C. (1998). From 'switching code' to 'code-switching': Towards a reconceptualization of communicative codes. In P. Auer (Ed.), *Code-switching in Conversation* (pp. 29-48). London: Routledge.
- Anderson, B. (2006 [1983]). *Imagined Communities: Reflections on the origin and spread of nationalism* (3rd ed.). London: Verso.
- ap Dyfrig, R., Jones, E. H. G., & Jones, G. (2006). The Welsh Language in the Media. *Mercator Media Monographs*, 1.
- Auer, P. (1995). The pragmatics of code-switching: a sequential approach. In L. Milroy & P. Muysken (Eds.), *One Speaker, Two Languages: Cross-disciplinary perspectives on code-switching* (pp. 115-135). Cambridge: Cambridge University Press.
- Auer, P. (2007). The monolingual bias in bilingualism research - or: Why bilingual talk is (still) a challenge for linguistics. In M. Heller (Ed.), *Bilingualism: A social approach* (pp. 319-339). Basingstoke; New York: Palgrave Macmillan.
- Bakhtin, M. (1935 [1981]). Discourse in the novel. In M. Holquist (Ed.), *The Dialogic Imagination: Four essays by M. M. Bakhtin* (pp. 259-422). Austin: University of Texas Press.
- Ball, M. J., Griffiths, T., & Jones, G. E. (1988). Broadcast Welsh. In M. J. Ball (Ed.), *The Use of Welsh: A contribution to sociolinguistics* (pp. 182-199). Clevedon: Multilingual Matters.
- Bevan, G. A. (2008). *Ailbeintio Pont y Borth? Y Geiriadur a'r Gymraeg*: Canolfan Uwchefrydiau Cymreig a Cheltaidd Prifysgol Cymru.
- Bhreathnach, Ú. (2009). *Is there a best-practice model for term planning?* Paper

- presented at the Language Planning and Language Learning conference, Limerick, Ireland, 19 June 2009.
- Bhreathnach, Ú. (2012). Term planning models and the Catalan case. *Revista de Llengua i Dret*, 58, 92-110.
- Bishop, H., Coupland, N., & Garrett, P. (2005). Conceptual accent evaluation: Thirty years of accent prejudice in the UK. *Acta Linguistica Hafniensia*, 37, 131-154.
- Blom, J.-P., & Gumperz, J. J. (1972). Social meaning in linguistic structures: Code-switching in Norway. In J. J. Gumperz & D. Hymes (Eds.), *Directions in Sociolinguistics* (pp. 407-434). New York: Holt, Rinehart and Winston.
- Bourdieu, P. (1991). *Language and Symbolic Power*. Cambridge: Polity Press.
- Brown, G. (1995). *Speakers, Listeners and Communication*. Cambridge: Cambridge University Press.
- Brown, P. (2000). Repetition. *Journal of Linguistic Anthropology*, 9(1-2), 223-226.
- Busch, B. (2010). New national languages in eastern Europe. In N. Coupland (Ed.), *The Handbook of Language and Globalisation* (pp. 182-200). Oxford: Wiley-Blackwell.
- Cabré Castellví, M. T. (2003). Theories of terminology: Their description, prescription and explanation. *Terminology*, 9(2), 163-199.
- Cameron, D. (2009 [1990]). Demythologizing sociolinguistics. In N. Coupland & A. Jaworski (Eds.), *The New Sociolinguistics Reader* (pp. 106-118). Basingstoke: Palgrave Macmillan.
- Carter, R. (1987). *Vocabulary: Applied linguistic perspectives*. London: Allen and Unwin.
- Cartrite, B. (2009). Minority language policy in France: Jacobinism, cultural pluralism, and ethnoregional identities. In M. H. Ross (Ed.), *Culture and Belonging in Divided Societies: Contestation and symbolic landscapes* (pp. 128-150). Philadelphia: University of Pennsylvania Press.
- Clyne, M. (1997). Introduction: On the undoing and redoing of corpus planning. In M. Clyne (Ed.), *Undoing and Redoing Corpus Planning* (pp. 1-9). Berlin; New York: Mouton de Gruyter.

- Colomer, R., & Cuadrado, S. (2007). *Catalan Terminological Planning: A model based on social participation*. Paper presented at the Minority Languages and Terminology Policies conference, Dublin, Ireland, 27-28 July 2007.
- Cooper, R. L. (1989). *Language Planning and Social Change*. Cambridge: Cambridge University Press.
- Cormack, M. (2007). Introduction: Studying minority language media. In M. Cormack & N. Hourigan (Eds.), *Minority Language Media: Concepts, critiques and case studies* (pp. 1-16). Clevedon: Multilingual Matters.
- Coupland, N. (2000). Sociolinguistic prevarication about 'standard English'. *Journal of Sociolinguistics*, 4(4), 622-634.
- Coupland, N. (2001). Language, situation and the relational self: theorizing dialect-style in sociolinguistics. In P. Eckert & J. R. Rickford (Eds.), *Style and Sociolinguistic Variation* (pp. 185-210). Cambridge: Cambridge University Press.
- Coupland, N. (2007). *Style: Language variation and identity*. Cambridge: Cambridge University Press.
- Coupland, N. (2009). The Mediated Performance of Vernaculars. *Journal of English Linguistics*, 37, 284-300.
- Coupland, N., & Kristiansen, T. (2011). SLICE: Critical perspectives on language (de)standardisation. In T. Kristiansen & N. Coupland (Eds.), *Standard Languages and Language Standards in a Changing Europe* (pp. 11-35). Oslo: Novus Press.
- David, R. (2013). Living Welsh in a globalised age. *Click on Wales*. Retrieved 18/08/2013, from <http://www.clickonwales.org/2013/07/living-welsh-in-a-globalised-world/>
- Davies, C. (1988). Cymraeg Byw. In M. J. Ball (Ed.), *The Use of Welsh: A contribution to sociolinguistics* (pp. 200-210). Clevedon: Multilingual Matters.
- Davies, D. (2000). *Arwyddocâd yr Iaith Gymraeg yn y Gwasanaeth Gofal / The Significance of the Welsh Language and Culture in the Health Care Services*. Paper presented at the Siarad yr Anweledig: Diwylliant, hunaniaeth a seiciatreg / Speaking the Invisible: Culture, identity and psychiatry conference, Felinfach, Wales, 4 November 2000.

- Davies, E. (1999). The Language of a Caring Service: Guidance on providing a sensitive bilingual service, focussing on the health, social care and justice sectors. Cardiff: Welsh Language Board.
- Davies, J. (2007). *Hanes Cymru*. London: Penguin.
- de Bres, J. (2008). Planning for tolerability in New Zealand, Wales and Catalonia. *Current Issues in Language Planning*, 9(4), 464-482.
- de Rooij, V. (1998). Problems of (Re-)contextualizing and interpreting variation in oral and written Shaba Swahili. *Cahiers de l'ILSL*, 10, 105-126.
- Drake, R. L., Vogl, W., & Mitchell, A. W. M. (2012). *Gray's Basic Anatomy*. Philadelphia: Elsevier Churchill Livingstone.
- Drakeford, M., & Morris, S. (1998). Social work with linguistic minorities. In C. Williams, H. Soydan & M. R. D. Johnson (Eds.), *Social Work and Minorities: European perspectives* (pp. 93-109). London: Routledge.
- Drame, A. (2007). *Terminology Policies and Minority Language Planning*. Paper presented at the Minority Languages and Terminology Policies conference, Dublin, 27-28 July 2007.
- Drame, A. (2008). Terminologieplanung ist weit mehr als Sprachplanung. *eDITion*, 2008 (1), 5-9.
- Duranti, A. (1997). *Linguistic Anthropology*. Cambridge: Cambridge University Press.
- Eckert, P., & McConnell-Ginet, S. (1992). Think practically and look locally: Language and gender as community-based practice. *Annual Review of Anthropology*, 21, 461-490.
- Eichmann, H. (2009). Planning sign languages: Promoting hearing hegemony? Conceptualizing sign language standardisation. *Current Issues in Language Planning*, 10(3), 293-307.
- Evans, W. G. (2000). Y wladwriaeth Brydeinig ac addysg Gymraeg 1914-1991. In G. H. Jenkins & M. A. Williams (Eds.), *'Eu Hiaith a Gadwant'? Y Gymraeg yn yr ugeinfed ganrif* (pp. 331-356). Cardiff: University of Wales Press.
- Fairclough, N. (1995a). *Critical Discourse Analysis: The critical study of language*. London; New York: Longman.

- Fairclough, N. (1995b). *Media Discourse*. London: Edward Arnold.
- Ffrancon, G., Dafydd, E., Prys, D., & Davies, O. (2008). Geiriadur Termau'r Diwydiannau Creadigol / Dictionary of Terms for the Creative Industries: Canolfan Addysg Uwch Cyfrwng Cymraeg.
- Fishman, J. A. (1983). Modeling rationales in corpus planning: Modernity and tradition in images of the good corpus. In J. Cobarrubias & J. A. Fishman (Eds.), *Progress in Language Planning: International perspectives* (pp. 107-118). Berlin; New York; Amsterdam: Mouton Publishers.
- Fishman, J. A. (1991). *Reversing Language Shift*. Clevedon: Multilingual Matters.
- Fishman, J. A. (2006). *Do Not Leave Your Language Alone: The hidden status agendas within corpus planning in language policy*. Mahwah, New Jersey; London: Lawrence Erlbaum Associates.
- Fishman, J. A. (2007 [1965]). Who speaks what language to whom and when? In Li Wei (Ed.), *The Bilingualism Reader* (2nd ed.). Abingdon; New York: Routledge.
- Gal, S., & Irvine, J. T. (1995). The boundaries of languages and disciplines: How ideologies construct difference. *Social Research*, 62(4), 967-1001.
- Gardner-Chloros, P. (2009). *Code-switching*. Cambridge: Cambridge University Press.
- Gaudin, F. (2003). *Socioterminologie: une approche sociolinguistique de la terminologie*. Bruxelles: Duculot.
- Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In H. Giles, J. Coupland & N. Coupland (Eds.), *Contexts of Accommodation* (pp. 1-68). Cambridge: Cambridge University Press.
- Goodwin, C. (1994). Professional Vision. *American Anthropologist*, 96(3), 606-633.
- Grace, G. (1990). The 'aberrant' (vs. 'exemplary') Melanesian languages. In P. Baldi (Ed.), *Linguistic Change and Reconstruction Methodology* (pp. 155-173). Berlin: Mouton de Gruyter.
- Griffiths, B., & Jones, D. G. (Eds.). (1995). *Geiriadur yr Academi / The Welsh Academy English-Welsh Dictionary*. Cardiff: University of Wales Press.
- Griffiths, D. C. (1993). *Termau Adeiladu*. Llangefni: Canolfan Astudiaethau Iaith.

- Gumperz, J. J. (1999). On interactional sociolinguistic method. In S. Sarangi & C. Roberts (Eds.), *Talk, Work and Institutional Order: Discourse in medical, mediation and management settings* (pp. 453-471). Berlin; New York: Mouton de Gruyter.
- Gwynedd Council. (2010). Cynllun Iaith Gymraeg / Welsh Language Scheme. Caernarfon.
- Haarmann, H. (1990). Language planning in the light of a general theory of language: A methodological framework. *International Journal of the Sociology of Language*, 86, 103-126.
- Haugen, E. (1953). *The Norwegian Language in America: A study in bilingual behavior. Vol. I: The Bilingual Community; Vol II: The American Dialects of Norwegian*. Bloomington: Indiana University Press.
- Haugen, E. (1966). Dialect, language, nation. *American Anthropologist*, 68(4), 922-935.
- Haugen, E. (1972 [1950]). The analysis of linguistic borrowing. In *The Ecology of Language* (pp. 79-109). Stanford: Stanford University Press.
- Haugen, E. (1972 [1961]). *Language Planning in Modern Norway*. Stanford: Stanford University Press.
- Haugen, E. (1983). The implementation of corpus planning: Theory and practice. In J. Cobarrubias & J. A. Fishman (Eds.), *Progress in Language Planning: International perspectives* (pp. 269-290). Berlin; New York; Amsterdam: Mouton Publishers.
- Heller, M. (2007). Bilingualism as ideology and practice. In M. Heller (Ed.), *Bilingualism: A social approach* (pp. 1-22). Basingstoke; New York: Palgrave Macmillan.
- Hicks, B., Morgan, M., Ó'Dochartaigh, C., Ellis, N., & Laporte, N. (1997). *Cysgair*. Bangor: Canolfan Bedwyr.
- Hill, J. (1985). The grammar of consciousness and the consciousness of grammar. *American Ethnologist*, 12(4), 725-737.
- Hill, J. (2000). Syncretism. *Journal of Linguistic Anthropology*, 9(1-2), 244-246.
- Hodges, R. (2009). Welsh language use among young people in the Rhymney valley.

- Contemporary Wales*, 22, 16-35.
- Holmes, J., & Meyerhoff, M. (1999). The community of practice: Theories and methodologies in language and gender research. *Language in Society*, 28, 173-183.
- Hornsby, M. (2009). *Globalisation Processes and Minority Languages: Linguistic hybridity in Brittany*. Unpublished PhD thesis, University of Southampton, Southampton.
- Hymes, D. (1972). Models of the interactions of language in social life. In J. J. Gumperz & D. Hymes (Eds.), *Directions in Sociolinguistics* (pp. 35-71). Oxford; New York: Basil Blackwell.
- Ignatieff, M. (1994). *Blood and Belonging: Journeys into the new nationalism*. London: Vintage.
- Information Commissioner's Office. (n.d.). Data Protection Act 1998: Legal Guidance. Retrieved 21/06/2013, from http://www.ico.gov.uk/upload/documents/library/data_protection/detailed_specialist_guides/data_protection_act_legal_guidance.pdf
- Iorwerth, R., James, E., & Williams, L. (2007). Termau Canŵio / Canoeing Terms: Welsh Language Board.
- Jaffe, A. (2003). Misrecognition unmasked?: 'Polynomic' language, expert statuses and orthographic practices in Corsican schools. *Pragmatics*, 13(4), 515-537.
- Jenkins, G. H. (1997). Adfywiad yr iaith a'r diwylliant Cymraeg 1660-1800. In G. H. Jenkins (Ed.), *Y Gymraeg yn ei Disgleirdeb: Yr iaith Gymraeg cyn y chwyldro diwydiannol* (pp. 365-400). Cardiff: University of Wales Press.
- Jernudd, B. H., & Das Gupta, J. (1971). Towards a theory of language planning. In J. Rubin & B. H. Jernudd (Eds.), *Can Language be Planned?* (pp. 195-216). Honolulu: University Press of Hawaii.
- Jilg, T. (2003). *Cadarnle'r Gymraeg? Astudiaeth Eirfaol o'r iaith Gymraeg ym Mlaenau Ffestiniog*. Unpublished PhD thesis, Cardiff University.
- Johnson, C., Dowd, T. J., & Ridgeway, C. L. (2006). Legitimacy as a social process. *Annual Review of Sociology*, 32, 53-78.

- Johnstone et al., B. (1994)¹⁶². Repetition in discourse: A dialogue. In B. Johnstone (Ed.), *Repetition in Discourse: Interdisciplinary perspectives* (Vol. I, pp. 1-20). Norwood, New Jersey: Ablex Publishing Corporation.
- Jones, A. E. (1982). Erydiad geirfaol: Astudiaeth ragarweiniol. *Papurau Gwaith Ieithyddol Cymraeg Caerdydd*, 2, 25-39.
- Jones, A. E. (1985). Erydiad geirfaol ym mhentrefi Clunderwen, Efailwen a Llandysilio. *Papurau Gwaith Ieithyddol Cymraeg Caerdydd*, 4, 81-98.
- Jones, B. P. (1988). Official Welsh. In M. J. Ball (Ed.), *The Use of Welsh: A contribution to sociolinguistics* (pp. 172-181). Clevedon: Multilingual Matters.
- Jones, D. G. (1988). Literary Welsh. In M. J. Ball (Ed.), *The Use of Welsh: A contribution to sociolinguistics* (pp. 125-171). Clevedon: Multilingual Matters.
- Jones, D. G. (2013). Welsh is a language just like any other. *Click on Wales* Retrieved 8/8/13, from <http://www.clickonwales.org/2013/08/welsh-is-a-language-just-like-any-other/>
- Jones, E. L. (2002). Buarth yr ysgol. *Barn*, 478, 4-7.
- Jones, M. C. (1998). *Language Obsolescence and Revitalization: Linguistic change in two sociolinguistically contrasting Welsh communities*. Oxford: Clarendon Press.
- Jones, T. (1688). Y Gymraeg yn ei Disgleirdeb, neu Helaeth Eir-lyfr Cymraeg a Saesneg / The British Language in its Lustre, or a Copious Dictionary of Welsh and English. London.
- Kaplan, R. B., & Baldauf, R. B., Jr. (1997). *Language Planning from Practice to Theory*. Clevedon: Multilingual Matters.
- Kiesling, S. F. (2004). Dude. *American Speech*, 79(3), 281-305.
- Kloss, H. (1969). Research possibilities on group bilingualism: A report. Quebec: International Centre for Research on Bilingualism.
- Lewis, D. G. (1996). *Termau Llywodraeth Leol / A Glossary of Local Government Terms*. Llandysul: Gomer.

¹⁶² Johnstone refers to her co-authors as all the participants who took part in discussions at the conference that led to the publication of the volume. I have not listed them all here for their sheer number.

- Lewis, Glyn (1983). Implementation of language planning in the Soviet Union. In J. Cobarrubias & J. A. Fishman (Eds.), *Progress in Language Planning: International perspectives* (pp. 309-326). Berlin; New York; Amsterdam: Mouton Publishers.
- Lewis, Gwion (2008). *Hawl i'r Gymraeg*. Talybont: Y Lolfa.
- Lewis, Glyn E. (1987). Attitude to the planned development of Welsh. *International Journal of the Sociology of Language*, 66, 11-26.
- Lewis, H. (1943). *Yr Elfen Ladin yn yr Iaith Gymraeg*. Cardiff: University of Wales Press.
- Léwis, R. (2003). *Geiriadur Newydd y Gyfraith / The New Legal Dictionary*. Llandysul: Gomer.
- Li Wei. (2000). Methodological issues in the study of bilingualism. In Li Wei (Ed.), *The Bilingualism Reader* (pp. 495-504). Abingdon; New York: Routledge.
- Li Wei. (2002). 'What do you want me to say?' On the Conversation Analysis approach to bilingual interaction. *Language in Society*, 31, 159-180.
- Liddicoat, A. J. (2000). The ecological impact of a dictionary. *Current Issues in Language Planning*, 1(3), 424-430.
- Lindsay, C. F. (1993). Welsh and English in the city of Bangor: A study in functional differentiation. *Language in Society*, 22, 1-17.
- Lippi-Green, R. (1997). *English with an Accent: Language, ideology, and discrimination in the United States*. London; New York: Routledge.
- Lloyd-Morgan, C. (2010). What's Welsh for woman? *Planet: The Welsh Internationalist*, 200, 26-31.
- Lyons, J. (1995). *Linguistic Semantics: An introduction*. Cambridge: Cambridge University Press.
- Mäkitalo, Å., & Säljö, R. (2002). Talk in interactional context and institutional context in talk: Categories as situated practices. *Text*, 22(1), 57-82.
- Martin-Jones, M. (1995). Code-switching in the classroom: Two decades of research. In L. Milroy & P. Muysken (Eds.), *One Speaker, Two Languages: Cross-disciplinary perspectives on code-switching* (pp. 90-111). Cambridge:

Cambridge University Press.

- May, S. (2000). Uncommon languages: The challenges and possibilities of minority language rights. *Journal of Multilingual and Multicultural Development*, 21(5), 366-385.
- Mayr, A. (2008). Introduction: Power, discourse and institutions. In A. Mayr (Ed.), *Language and Power* (pp. 1-25). London; New York: Continuum.
- McCarty, T. L. (2011). Introducing ethnography and language policy. In T. L. McCarty (Ed.), *Ethnography and Language Policy* (pp. 1-28). New York; London: Routledge.
- McEnery, T., & Wilson, A. (2001). *Corpus Linguistics* (2nd ed.). Edinburgh: Edinburgh University Press.
- MEU Cymru. (1992). Geiriadur Termau Cyfrifiadureg / Dictionary of Computer Terms. Cardiff: MEU Cymru.
- Milroy, J. (1999). The consequences of standardisation in descriptive linguistics. In T. Bex & R. J. Watts (Eds.), *Standard English: The widening debate* (pp. 16-39). London; New York: Routledge.
- Milroy, J. (2001). Language ideologies and the consequences of standardization. *Journal of Sociolinguistics*, 5(4), 530-555.
- Milroy, J. (2002). The legitimate language: Giving a history to English. In R. Watts & P. Trudgill (Eds.), *Alternative Histories of English* (pp. 7-25). London; New York: Routledge.
- Milroy, J., & Milroy, L. (1997). Varieties and variation. In F. Coulmas (Ed.), *The Handbook of Sociolinguistics* (pp. 47-64). Oxford: Blackwell.
- Milroy, J., & Milroy, L. (2012 [1985]). *Authority in Language: Investigating Standard English* (4th ed.). London; New York: Routledge.
- Milroy, L. (1999). Standard English and language ideology in Britain and the United States. In T. Bex & R. J. Watts (Eds.), *Standard English: The widening debate* (pp. 173-206). London: Routledge.
- Misell, A. (2000). Y Gymraeg yn y Gwasanaeth Iechyd: Ehangder, natur a digonolrwydd darpariaeth Gymraeg yn y Gwasanaeth Iechyd Gwladol yng Nghymru / Welsh in the Health Service: The scope, nature and adequacy of

- Welsh Language Provision in the National Health Service in Wales: Consumer Council for Wales.
- Mishler, E. G. (1984). *The Discourse of Medicine: Dialectics of medical interviews*. Norwood: Ablex Publishing Corporation.
- Moon, R. (1997). Vocabulary connections: Multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 40-63). Cambridge: Cambridge University Press.
- Morgan, D. D. (2000). Yr Iaith Gymraeg a chrefydd. In G. H. Jenkins & M. A. Williams (Eds.), *'Eu Hiaith a Gadwant'? Y Gymraeg yn yr ugeinfed ganrif* (pp. 357-380). Cardiff: University of Wales Press.
- Morgan, M. (2002). *Hanes geiriaduraeth yng Nghymru o 1547 hyd 1914: gyda sylw arbennig i ddylanwad John Walters a William Owen Pughe ar eiriadurwyr 1805-1850*. Unpublished PhD thesis, University of Wales.
- Morris, D. (1989). A study of language contact and social networks in Ynys Môn. *Contemporary Wales*, 3, 99-117.
- Mugglestone, L. (2011). *Dictionaries: A very short introduction*. Oxford: Oxford University Press.
- Mwansoko, H. J. M. (1993). Acceptability of terms: The case of the Swahili linguistics and literature terms in Tanzania. *Multilingua*, 12(2), 177-188.
- Myers-Scotton, C. (1993). *Social Motivations for Code-switching*. Oxford: Clarendon Press.
- Ní Ghearáin, H. (2008). Irish terminology planning. *eDITion*, 14-17.
- Ní Ghearáin, H. (2009). *Terminology planning in the lesser-used language and the dynamics of acceptance: insights from sociolinguistic research in the Gaeltacht*. Paper presented at the Language Planning and Language Learning conference, Limerick, Ireland, 19 June 2009.
- Oakes, L. (2001). *Language and National Identity: Comparing France and Sweden*. Amsterdam; Philadelphia: John Benjamins.
- Ochs, E. (2009). Linguistic resources for socializing humanity. In N. Coupland & A. Jaworski (Eds.), *The New Sociolinguistics Reader* (pp. 405-423). Basingstoke: Palgrave Macmillan.

- Oxford University Press. (2013). Oxford English Dictionary online. Retrieved at various dates between 2008 and 2013, from www.oed.com
- Pennycook, A. (2010). *Language as a Local Practice*. Abingdon: Routledge.
- Phillips, D. (2000). Hanes Cymdeithas yr Iaith Gymraeg 1962-1998. In G. H. Jenkins & M. A. Williams (Eds.), *'Eu Hiaith a Gadwant?': Y Gymraeg yn yr ugeinfed ganrif* (pp. 445-472). Cardiff: University of Wales Press.
- Poplack, S., Sankoff, D., & Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, 26, 47-104.
- Porter, R. (1997). *The Greatest Benefit to Mankind: A medical history of humanity from antiquity to the present*. London: Fontana Press.
- Pryce, W. T. R., & Williams, C. H. (1988). Sources and methods in the study of language areas: A case study of Wales. In C. H. Williams (Ed.), *Language in Geographic Context* (pp. 167-237). Clevedon: Multilingual Matters.
- Prys, D. (1998). Termau Deddfwriaeth Priffyrdd / Highways Legislation Terminology. Cardiff: Welsh Language Board.
- Prys, D. (2000a). Termau Cyllid y Cynulliad / Dictionary of Finance Terms for the National Assembly of Wales. Cardiff: Welsh Language Board.
- Prys, D. (2000b). Termau Gwaith a Gofal Cymdeithasol / Social Work and Social Care Terms. Cardiff: CCETSW.
- Prys, D. (2000c). Termau Hybu Iechyd / Terms for Health Promotion. Bangor: University of Wales, Bangor.
- Prys, D. (2002a). Termau Asiantaeth yr Amgylchedd Cymru / Environment Agency Wales Terms. Bangor: University of Wales Bangor / Environment Agency Wales.
- Prys, D. (2002b). Termau Iechyd Meddwl Plant a Phobl Ifanc / Child and Adolescent Mental Health Terms. Bangor: University of Wales, Bangor.
- Prys, D. (2005). Termau Gofal Iechyd Pobl Hŷn / Terms in the Health Care of Older People. Bangor: University of Wales, Bangor.
- Prys, D. (2006). Setting the Standards: Ten years of Welsh terminology work. In P. Ten Hacken (Ed.), *Terminology, Computing and Translation* (pp. 41-55).

Tübingen: Gunter Narr Verlag.

- Prys, D. (2007a). *Cyflwyniad ar Grŵp Safoni Termau Gofal Iechyd Gogledd Cymru*. Paper presented at the Seminar Termau Bwrdd yr Iaith Gymraeg / Welsh Language Board Terminology Seminar, 30 October 2007.
- Prys, D. (2007b). *Terminology Policies in Wales*. Paper presented at the Minority Languages and Terminology Policies conference, Dublin, 27-28 July 2007.
- Prys, D., Davies, O., & Prys, G. (2005). *Termau Technoleg Gwybodaeth / Information Technology Terms*. Cardiff: Welsh Language Board.
- Prys, D., Davies, O., & Prys, G. (2006). *Rhestr Arwyddion Siopau / Shop Signs Glossary*: Welsh Language Board.
- Prys, D., & Davies, O. L. (2007). *Termau Therapi Galwedigaethol / Terms for Occupational Therapy*. Bangor: University of Wales, Bangor.
- Prys, D., & Jones, D. B. (2007). *Canllawiau Safoni Termau Gwasanaeth Cyfieithu Llywodraeth Cynulliad Cymru a Bwrdd yr Iaith Gymraeg / Guidelines for the Standardization of Terminology for the Welsh Assembly Government and the Welsh Language Board: Welsh Assembly Government / Welsh Language Board*.
- Prys, D., & Jones, J. P. M. (1998a). *Report on the Standardization of Terms Project*: Welsh Language Board.
- Prys, D., & Jones, J. P. M. (1998b). *Y Termiadur Ysgol: Termau wedi'u safoni ar gyfer ysgolion Cymru*. Cardiff: Qualifications, Curriculum and Assessment Authority for Wales.
- Prys, D., Jones, J. P. M., Davies, O., & Prys, G. (2006). *Y Termiadur*. Cardiff: Qualifications, Curriculum and Assessment Authority for Wales.
- Prys, D., Jones, J. P. M., & Emlyn, H. a. (1995). *Llyfryddiaeth Geiriaduron Termau*. Bangor: University of Wales Press.
- Prys, D., & Williams, L. (2001). *Anabledd ac Iaith: Canllawiau ar gyfer defnyddio terminoleg anabledd / Disability and Language: Guidelines for the use of disability terms: Anabledd Cymru / Disability Wales*.
- Pughe, W. O. (1793-1803). *A Dictionary of the Welsh Language, explained in English: with numerous illustrations, from the literary remains and from the living*

- speech of the Cymry. London.
- Quirion, J. (2003). Methodology for the design of a standard research protocol for measuring terminology usage. *Terminology*, 9(1), 29-49.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. Harlow: Longman.
- Rampton, B. (2007). Neo-Hymesian linguistic ethnography in the United Kingdom. *Journal of Sociolinguistics*, 11(5), 584-607.
- Robb, G. (2007). *The Discovery of France*. London: Picador.
- Robert, E. (2009). Accommodating "new" speakers? An attitudinal investigation of L2 speakers of Welsh in south-east Wales. *International Journal of the Sociology of Language*, 195, 93-116.
- Robert, E. (2011). Standardness and the Welsh language. In T. Kristiansen & N. Coupland (Eds.), *Standard Languages and Language Standards in a Changing Europe* (pp. 135-143). Oslo: Novus.
- Roberts, G. (2000). Bilingualism and number in Wales. *International Journal of Bilingual Education and Bilingualism*, 3(1), 44-56.
- Roberts, G., Irvine, F., Jones, P., Spencer, L., Baker, C., & Williams, C. (2004). *Adroddiad am Astudiaeth o Ymwybyddiaeth o'r Gymraeg yn y Ddarpariaeth Gofal Iechyd yng Nghymru*. Bangor: University of Wales, Bangor.
- Roberts, G., & Prys, D. (1995). *Termau Nyrsio a Bydwreigiaeth / An English-Welsh Dictionary of Nursing and Midwifery Terms*. Bangor: University of Wales, Bangor.
- Rock, F. (2001). Policy and practice in the anonymisation of linguistic data. *International Journal of Corpus Linguistics*, 6(1), 1-26.
- S4C. (2008). Canllawiau Iaith.
- S4C. (2010). S4C Programme Guidelines.
- S4C. (n.d. [c. 2001]). Canllawiau Cydymffurfiaeth / Compliance Guidelines.
- Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696-735.

- Salesbury, W. (1547). *A Dictionary in Englyshe and Welsh moche necessary to all such Welshemen, as wil spedly learne the englyshe to'gue*. London.
- Sarangi, S., & Clarke, A. (2002). Zones of expertise and the management of uncertainty in genetics risk communication. *Research on Language and Social Interaction*, 35(2), 139-171.
- Schiffman, H. F. (1992). 'Resisting arrest' in status planning: Structural and covert impediments to status change. *Language & Communication*, 12(1), 1-15.
- Schiffman, H. F. (1996). *Linguistic Culture and Language Policy*. London: Routledge.
- Schiffrin, D. (2006). *In Other Words: Variation in reference and narrative*. Cambridge: Cambridge University Press.
- Shapiro, M. J. (1989). A political approach to language purism. In B. H. Jernudd & M. J. Shapiro (Eds.), *The Politics of Language Purism* (pp. 21-29). Berlin; New York: Mouton de Gruyter.
- Shohamy, E. (2006). *Language Policy: Hidden agendas and new approaches*. London; New York: Routledge.
- Smith, R. (2000). Darlledu a'r iaith Gymraeg. In G. H. Jenkins & M. A. Williams (Eds.), *'Eu Hiaith a Gadwant'?: Y Gymraeg yn yr ugeinfed ganrif* (pp. 299-329). Cardiff: University of Wales Press.
- Spencer, L., Edwards, M., Prys, D., & Thomas, E. (2004). *Geiriadur Termiau Seicoleg / Dictionary of Terms for Psychology*. Bangor: Bangor University.
- Spolsky, B. (2004). *Language Policy*. Cambridge: Cambridge University Press.
- Stammers, J. R. (2009). *The Integration of English-origin Verbs in Welsh*. Unpublished PhD thesis, Bangor University.
- Stavans, I. (2003). *Spanglish: The making of a new American language*. New York: Harper Collins.
- Stroud, C. (1992). The problem of intention and meaning in code-switching. *Text*, 12(1), 127-155.
- Stubbe, M., Lane, C., Hilder, J., Vine, E., Vine, B., Marra, M., et al. (2003). Multiple discourse analyses of a workplace interaction. *Discourse Studies*, 5, 351-388.

- Temmerman, R. (1997). Questioning the univocity ideal: The difference between socio-cognitive terminology and traditional terminology. *Hermes, Journal of Linguistics*, 18, 51-90.
- The British Association for Applied Linguistics. (2006). Recommendations on Good Practice in Applied Linguistics. Retrieved 23/11/08, from www.baal.org.uk/about_goodpractice_full.pdf
- Thomas, A. R. (1982). Change and decay in language. In D. Crystal (Ed.), *Linguistic Controversies: Essays in linguistic theory and practice in honour of F. R. Palmer* (pp. 209-219). London: Edward Arnold.
- Thomas, A. R. (1987). A spoken standard for Welsh: Description and pedagogy. *International Journal of the Sociology of Language*, 66, 99-113.
- Thomas, B., & Thomas, P. W. (1989). *Cymraeg, Cymrâg, Cymrêg... Cyflwyno'r tafodieithoedd*. Cardiff: Gwasg Taf.
- Thomas, C. H. (1967). Review of *Cymraeg Byw, Rhifyn 2*. Welsh Joint Education Committee, 1967. *Llên Cymru*, 13, 242-249.
- Thomas, C. H. (1982). Registers in Welsh. *International Journal of the Sociology of Language*, 35, 87-115.
- Thomas, G. (1991). *Linguistic Purism*. London; New York: Longman.
- Thurlow, C., & Poff, M. (2013). Text messaging. In S. C. Herring, D. Stein & T. Virtanen (Eds.), *Pragmatics of Computer-Mediated Communication* (pp. 163-190). Berlin: Mouton de Gruyter.
- Tollefson, J. W. (2006). Critical theory in language policy. In T. Ricento (Ed.), *An Introduction to Language Policy: Theory and method* (pp. 42-59). Malden: Blackwell Publishing.
- University of Wales Press. (1950-2002). *Geiriadur Prifysgol Cymru*. Retrieved 25/02/2009, 2009, from www.cymru.ac.uk/geiriadur
- University of Wales Press. (1957). *Termau Mathemateg*. Cardiff: University of Wales Press.
- University of Wales Press. (1959). *Termau Hanes*. Cardiff: University of Wales Press.
- University of Wales Press. (1965). *Termau Ffiseg a Mathemateg*. Cardiff: University

- of Wales Press.
- University of Wales Press. (1972). *Termau Economeg ac Econometreg*. Cardiff: University of Wales Press.
- University of Wales Press. (1978). *Termau Cerddoriaeth*. Cardiff: University of Wales Press.
- University of Wales Press. (1986). *Termau Meddygol / Medical Terms*: University of Wales Press.
- Verschueren, J. (2004). Notes on the role of metapragmatic awareness in language use. In A. Jaworski, N. Coupland & D. Galasiński (Eds.), *Metalinguage: Social and ideological perspectives* (pp. 53-73). Berlin; New York: Mouton de Gruyter.
- Walters, J. (1794). *An English-Welsh Dictionary, wherein not only the words, but also, the idioms and phraseology of the English language, are carefully translated into Welsh, by proper and equivalent words and phrases*. London.
- Weinstein, B. (1989). Francophonie: Purism at the international level. In B. H. Jernudd & M. J. Shapiro (Eds.), *The Politics of Language Purism* (pp. 53-79). Berlin; New York: Mouton de Gruyter.
- Wells, G. (2007). Semiotic mediation, dialogue and the construction of knowledge. *Human Development*, 50, 244-274.
- Welsh Assembly Government. (2003). *Iaith Pawb: A national action plan for a bilingual Wales*. Cardiff: Welsh Assembly Government.
- Welsh Government. (2012a). *2011 Census: First results on the Welsh Language*. Cardiff.
- Welsh Government. (2012b). *Iaith Fyw: Iaith Byw / A Living Language: A Language for Living*. Cardiff.
- Welsh Government. (2013). *A Living Language: A Language for Living: Welsh Language Strategy 2012-17: Action Plan 2013-14*. Cardiff.
- Welsh Joint Education Committee. (1967). *Cymraeg Byw* (Vol. 2). Cardiff: Welsh Joint Education Committee.
- Welsh Joint Education Committee. (1970). *Cymraeg Byw* (Vol. 3). Cardiff: Welsh Joint

- Education Committee.
- Welsh Joint Education Committee. (1983). *Termau Ffiseg a Mathemateg*. Cardiff: Welsh Joint Education Committee.
- Welsh Joint Education Committee. (1986). *Termau Cyfrifiadureg*. Cardiff: Welsh Joint Education Committee.
- Welsh Joint Education Committee. (1987). *Termau Hanes*. Cardiff: Welsh Joint Education Committee.
- Welsh Joint Education Committee. (1992a). *Termau Cerddoriaeth*. Cardiff: Welsh Joint Education Committee.
- Welsh Joint Education Committee. (1992b). *Termau Economeg*. Cardiff: Welsh Joint Education Committee.
- Welsh Language Board. (1995). *The Panel for Official Welsh: Report*. Cardiff: Welsh Language Board.
- Welsh Language Board. (2005). *The Future of Welsh: A Strategic Plan*. Cardiff: Welsh Language Board.
- Welsh Language Board. (2008). *Canolfan Safoni Termau Genedlaethol ar ei ffordd*. Retrieved 20/08/2013, from www.byig-wlb.org.uk/cymraeg/newyddion/pages/canolfansafonitermaugenedlaethola-reiffordd.aspx
- Welsh Language Board. (2009a). *Hyrwyddo a Hwyluso Gweithleoedd Dwyieithog / Promoting and Facilitating Bilingual Workplaces*. Cardiff.
- Welsh Language Board. (2009b). *Minutes of the Welsh Language Board's Translation Forum*. Cardiff: Welsh Language Board.
- Wexler, P. N. (1974). *Purism and Language: A study in modern Ukrainian and Belorussian Nationalism (1840-1967)*. Bloomington: Indiana University Press.
- Williams, C. (1999). *Cymraeg Clir*. Bangor: Gwynedd Council, Welsh Language Board, Canolfan Bedwyr.
- Williams, C. H. (2000). *Adfer yr iaith*. In G. H. Jenkins & M. A. Williams (Eds.), *'Eu Hiaith a Gadwant'? Y Gymraeg yn yr ugeinfed ganrif* (pp. 641-665). Cardiff: University of Wales Press.

- Williams, E. (2009). Language attitudes and identity in a north Wales town: "Something different about Caernarfon"? *International Journal of the Sociology of Language*, 195, 63-91.
- Williams, G. (1987). Bilingualism, class dialect, and social reproduction. *International Journal of the Sociology of Language*, 66, 85-98.
- Williams, G. (1992). *Sociolinguistics: A sociological critique*. London, New York: Routledge.
- Williams, G., & Thomas, A. R. (n.d.). *Welsh Language Output of S4C: A feasibility study aimed to develop techniques to study the relationship between the Welsh language output of S4C, and any changes to the extended pattern of use, and in the linguistic characteristics of the language over time*. Cardiff: The Welsh Office.
- Williams, I. (1960). Cymraeg Byw, *Annual radio lecture, 1 March 1960*: BBC Wales.
- Winford, D. (2003). *An Introduction to Contact Linguistics*. Malden: Blackwell.
- Woolard, K. A. (1985). Language variation and cultural hegemony: toward an integration of sociolinguistic and social theory. *American Ethnologist*, 12(4), 738-748.
- Woolard, K. A. (1998). Simultaneity and bivalency as strategies in bilingualism. *Journal of Linguistic Anthropology*, 8(1), 3-29.
- Woolard, K. A. (2005). Codeswitching. In A. Duranti (Ed.), *A Companion to Linguistic Anthropology*. Blackwell Reference Online: Blackwell Publishing.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
- Wright, S. (2004). *Language Policy and Language Planning: From Nationalism to Globalisation*. Houndmills: Palgrave Macmillan.
- Wright, S. (2007). The right to speak one's own language: Reflections on theory and practice. *Language Policy*, 6, 203-224.

APPENDIX ONE:

Information sheet and consent form

I include here the text from the information sheet and consent form for individuals at the workplace. I have anonymised the text, and omitted the formatting and layout.

Ymchwil iaith a chyfathrebu – Amlinelliad prosiect

Rydw i'n gwneud gwaith ymchwil ar iaith a chyfathrebu ac yn gwneud rhan o'r gwaith yn (organisation name). Rydw i'n casglu data ar gyfer yr astudiaeth yn rhai o gyfarfodydd (organisation name). Hoffwn i gael eich caniatâd chi i recordio cyfarfodydd byddwch chi'n rhan ohonynt gyda (organisation name). Er mwyn helpu i chi ystyried ydych chi am gymryd rhan, dyma wybodaeth am yr ymchwil:

Beth yw pwrpas y gwaith?

Rydw i'n edrych ar sut mae pobl yn cyfathrebu ac yn defnyddio iaith, yn benodol y geiriau a'r termau mae pobl yn eu defnyddio. Er enghraifft, ydy pobl yn defnyddio termau Cymraeg, ydy termau Cymraeg yn rhy ffurfiol neu ydyn nhw wedi dod yn rhan normal o iaith bob dydd? Bwriad y gwaith yw gweld beth mae pobl yn ei wneud ac yn ei feddwl am iaith, *nid* plismona iaith na phasio barn am allu ieithyddol.

Beth fydd yn digwydd os bydda i'n cytuno?

Os byddwch chi'n cytuno i gymryd rhan, bydda i'n recordio rhai cyfarfodydd byddwch chi'n rhan ohonynt. Bydda i wedyn yn trawsgrifio'r recordiadau. Mae hynny'n golygu teipio beth sy'n digwydd ac yn cael ei ddweud. Wedyn, bydda i'n dadansoddi'r trawsgrifiadau. Bydda i'n adrodd yn ôl i (organisation name) ac i ymchwilwyr eraill mewn prifysgolion. Bydda i hefyd yn cyflwyno fy ngwaith i fy noddwyr, Bwrdd yr Iaith Gymraeg.

Beth fydd yn digwydd os fydda i ddim yn cytuno?

Rydw i'n gobeithio y byddwch chi'n cytuno i gymryd rhan yn yr astudiaeth. Fodd bynnag, does dim rhaid i chi gymryd rhan. Os na fyddwch chi'n cytuno, fydd hynny ddim yn effeithio ar eich perthynas chi gyda (organisation name).

Sut caiff fy hawliau i eu gwarchod?

Mae yna rai amodau i'r ymchwil er mwyn sicrhau y caiff eich hawliau chi eu gwarchod.

- O ran y recordio ei hun:
 - gallwch chi ofyn am gael stopio'r recordio ar unrhyw adeg yn ystod cyfarfod
 - gallwch chi ofyn am gael dileu darn o'r recordiad ar ôl y cyfarfod

- gallwch chi ofyn am gael tynnu nôl o'r astudiaeth ar ôl recordio Nodwch na fydd modd i chi dynnu nôl o'r astudiaeth neu ofyn am gael dileu data ar ôl i mi wneud gwaith dadansoddi sylweddol nac ar ôl i mi gyflwyno canlyniadau. Felly, os byddwch chi'n penderfynu tynnu nôl neu ofyn am gael dileu data, mae'n bwysig eich bod chi'n gwneud hynny cyn pen mis i ddyddiad y cyfarfod.
- O ran y trawsgrifio, storio'r data a chyflwyno'r ymchwil:
 - wrth drawsgrifio, bydda i'n defnyddio ffugenwau ar gyfer pawb yn y recordiadau ac yn newid unrhyw wybodaeth arall sy'n digwydd codi a allai ddatgelu pwy ydych chi (e.e. enwau tai, cyfeiriadau, rhifau ffôn, enwau sefydliadau, teitlau swyddi)
 - bydda i'n cadw'r recordiadau a'r trawsgrifiadau ar ffurf ddigidol mewn ffeiliau dan gyfrinair
 - dim ond fi a fy nau diwtor ym Mhrifysgol Caerdydd fydd yn cael clywed y recordiadau'n llawn a darllen y trawsgrifiadau'n llawn
 - wrth gyflwyno fy ngwaith (mewn erthyglau, cyflwyniadau llafar ayb) bydda i'n gwneud yn siŵr na fydda i'n datgelu eich enw nac unrhyw wybodaeth arall a allai ddatgelu pwy ydych chi (fel uchod)

Cysylltu

Os oes gyda chi unrhyw sylwadau neu gwestiynau am yr ymchwil, baswn i'n falch iawn o glywed ganddoch chi, dros y ffôn, e-bost, neu drwy'r post, ac yn Gymraeg neu Saesneg:

Elen Robert, Canolfan Iaith a Chyfathrebu, Prifysgol Caerdydd, Adeilad y Dyniaethau, Colum Drive, Caerdydd CF10 3EU
(my email address and phone number)

Beth nesaf?

Os ydych chi'n hapus i gymryd rhan yn yr ymchwil, a wnewch chi lenwi'r manylion isod yn rhoi'ch caniatâd? Rhwygwch y darn yma oddi ar y daflen a chadw'r rhan arall, rhag ofn bydd gyda chi unrhyw gwestiynau eto.

Ymchwil iaith a chyfathrebu – Ffurflen caniatâd gan unigolion

Rydw i wedi darllen ac wedi deall yr amlinelliad o'r prosiect uchod, ac yn cytuno i gymryd rhan yn yr ymchwil. Rydw i'n rhoi caniatâd i gael fy recordio yng nghyfarfodydd (organisation name) ac i'r recordiadau gael eu trawsgrifio a'u dadansoddi gan yr ymchwilydd (Elen Robert), yn ôl yr amodau a ddisgrifir yn yr amlinelliad o'r prosiect. Rydw i'n deall os oes gen i unrhyw gwestiynau am y gwaith, galla i gysylltu ag Elen Robert ym Mhrifysgol Caerdydd, ac mae gen i gopi o'i manylion cysylltu.

Enw (priflythrennau) _____

Arwyddwyd _____

Dyddiad _____

Language and communication research – Project outline

I'm doing research on language and communication and carrying out some of the work at (organisation name). I'm collecting data for the study in some of (organisation name)'s meetings. I would like your consent to record some meetings you will be attending with (organisation name). To help you consider whether you want to take part, here is some information about the research:

What are the aims of the research?

I'm looking at how people communicate and use language, in particular the words and terms people use. For example, do people use Welsh terms, are Welsh terms too formal, or have they become a normal part of everyday language? The aim of the research is to see what people do and how people think about language, and *not* to pass judgement on language competence.

What will happen if I agree?

If you agree to take part, I will record some of the meetings you attend. I will then transcribe the recordings. This involves typing what happens and what is said. Then I will analyse the transcripts. I will report back to (organisation name) and to other university researchers. I will also present my work to my sponsors, the Welsh Language Board.

What will happen if I don't agree?

I hope you will agree to take part in the research. However, you don't have to take part. If you don't agree, it will not have any effect on your relationship with (organisation name).

How will my rights be protected?

There are some conditions to the research to make sure that your rights are protected.

- In terms of the actual recording:
 - you can ask to stop the recording at any time during a meeting
 - you can ask to delete a part of the recording after the meeting
 - you can ask to withdraw from the study after recording

Please note that you won't be able to withdraw from the study or ask for data to be deleted after I've done substantial analysis or after I present results.

Therefore, if you decide to withdraw or ask for data to be deleted, it's important that you do so within a month after the date of the meeting.

- In terms of transcribing, storing data and presenting the research:
 - in transcribing, I will use pseudonyms for everybody in the recordings and will change any other information that arises that could reveal your identity (e.g. house names, addresses, phone numbers, names of organisations, job titles)
 - I will keep the recordings and transcripts in digital format in password-protected files
 - only I and my two supervisors at Cardiff University will be able to listen to the recordings and read the transcripts in full
 - when I present my work (in articles, spoken presentations etc.) I will make sure that I don't reveal your name or any other information that could reveal your identity (as above)

Contact

If you have any comments or questions about the research, I'd be very glad to hear from you, by phone, e-mail or post, and in English or Welsh:

Elen Robert, Centre for Language and Communication, Cardiff University,
Humanities Building, Colum Drive, Cardiff CF10 3EU
(my email address and phone number)

What next?

If you are willing to take part in the research, please fill in the details below giving your consent. Tear off this part of the leaflet and keep the other part, in case you have any questions later.

Language and communication research – Consent form for individuals

I have read and understood the project outline above, and agree to take part in the research. I give consent to be recorded at meetings with (organisation name) and for the recordings to be transcribed and analysed by the researcher (Elen Robert), in keeping with the conditions described in the project outline. I understand that if I have any questions about the work, I can contact Elen Robert at Cardiff University, and I have a copy of her contact details.

Name (in capitals) _____

Signed _____

Date _____

APPENDIX TWO:

Sample transcript – *Syrjeri* data

Transcription conventions here are different to the ones I've used for the presentation of data excerpts in the body of the thesis. Specifically, here transcription of overlap and sequence is more rough than in the body of the thesis, where the presentation of sequence is potentially more important for analysis. Here overlap is marked by +< at the beginning of a line, which marks that the following line overlaps with part or all of the preceding line. Lines in this transcript are typically clause-based, due to the typically very long turns of Dr Ann's expositions.

Diabetes

Participants: Elinor Jones (EJ, presenter), Doctor Ann (DA), Guest (GU)

Situation: Studio chat between participants. DA is a weekly guest to the show, acting as medical expert. GU is a one-off guest who suffers from the condition, and who has come in to talk about his experiences of living with it.

Setting: Participants sit on two sofas that are at right angles to each other (like an upside-down 'V'). EJ sits on the right-hand sofa, DA and GU sit on the left-hand sofa. There is a screen on the wall behind them for displaying related images, but otherwise showing the Wedi3 logo. Camera shots vary from full studio shots to shots of individual participants. At times the broadcast cuts to still frames of various images, overlaid by audio from the studio. Information on visual data is transcribed below when considered relevant.

1. EJ: croeso nôl aton ni
welcome back
2. EJ: clefyd y siwgr (.) clefyd melys (.) neu (.) diabetes (.) dyna fyddwn ni'n ei (.)
drafod yn y syrjeri heddi
*diabetes (.) diabetes (.) or diabetes [gives three different names] (.) that's
what we'll be (.) discussing in the surgery today*
3. EJ: ac er fod na nifer fawr o bobl ym Mhrydain yn diodde o'r cyflwr (.) nid
pawb sy'n ymwybodol o hynny
*and although there are a large number of people in Britain who suffer from
the condition (.) not everybody is aware of that*

4. EJ: y ddau fath mwya cyffredin yw math un (.) a math dau
the two most common types are type one (.) and type two
5. EJ: ac er eu bod nhw'n cael eu hachosi gan ffactore gwahanol (.) yr un yw'r
canlyniad sef gormod o siwgr neu glwcos yn y gwaed
*and although they're caused by different factors (.) the result is the same,
namely too much sugar, or glucose, in the blood*
6. EJ: ac i drafod y cyflwr mae (.) (guest name) sy'n ddioddefwr a hefyd doctor
Ann
*and here to talk about the condition is (.) (guest name), who's a sufferer, and
also doctor Ann*
7. EJ: croeso i'r ddau noch chi
welcome both of you
8. EJ: dechreua i gyda chi Ann i ddechre xxx am y gwahanol fathe yma
I'll start with you Ann, to begin xxx about these different types
9. DA: ie
yes
10. EJ: ac beth sy'n digwydd yn y corff pan bod hwn yn digwydd i chi
and what happens in the body when this happens to you
11. DA: +< ie (.) ie
yes (.) yes
12. DA: yn y bôn fel wedoch chi yr un salwch yw'r ddau math un a math dau
basically, as you said, both types are the same illness, type one and type two
13. DA: oherwydd mae'r corff yn colli y gallu (.) i reoli (.) lefel y glwcos yn y gwaed
(.) oherwydd (.) methiant y pancreas
*because the body loses the ability (.) to control (.) the level of glucose in the
blood (.) because of (.) failure of the pancreas*
14. DA: mae gyda ni lun o'r pancreas i ni gael gwybod le mae'r organ ma
we've got a picture of the pancreas so we can know where this organ is
15. *(an image appears on the screen behind them and is then shown on full
screen. It consists of two colour diagrams, one in the top left-hand corner,
the other diagonally opposite. The diagram in the top left-hand corner shows
a woman's abdomen. It mostly shows the outer body, but shows some of the*

inner organs where they would be around the area of the stomach. The diagram in the bottom right-hand corner is a close-up of some of the inner organs shown in the other diagram. It labels 'Pancreas' and 'Duodenum'.)

16. DA: <mae fe> [?] fynna'n cuddio (.) dan yr afu dan y stumog
<it's> [?] there hiding (.) under the liver under the stomach

17. DA: a pwrpas un pwrpas pwysig y pancreas yw cynhyrchu inswlin
and the purpose...one important purpose of the pancreas is to produce insulin

18. DA: a mae'r llun nesa yn dangos celloedd arbennig yn y pancreas (.) sy yn cynhyrchu (.) yr inswlin
and the next picture shows special cells in the pancreas (.) that produce (.) the insulin

19. *(the image on the screen has changed to a visual showing several diagrams. On the left, vertically in the middle, is a small sketch of a male abdomen. It shows the outer body, but some internal organs around the area of the stomach. This organ looks very similar to what is represented in the larger diagram directly above it. In that diagram (in the top left-hand corner) there is a representation of the pancreas, labelled 'Pancreas'. To the right of it is a microscopic representation of a part of the previous diagram. Some purple blobs (that look something like purple frog spawn) are labelled 'Celloedd Insulin' (Insulin Cells). Below this diagram and slightly to the left is a microscopic representation of one of these purple blobs, labelled 'Cell Insulin' (Insulin Cell), and underneath the insulin cell (within the same diagram) is a blue arrow pointing into a pink channel. A disorderly row of brown flecks that are labelled 'Insulin' follow this arrow and the pink channel, across the diagram boundary and into the next diagram diagonally down to the left. In this diagram the pink channel continues, as do the brown flecks. There are also some grey-green flecks, labelled 'Glucose'. The grey-green and brown flecks follow a blue arrow into some red tubes up and to the left.)*

20. DA: nawr mae'r inswlin mae'r cell fynna ar y gwaelod tua chwech o'r gloch (.) yn cynhyrchu'r inswlin
now the insulin...the cell there on the bottom about six o'clock (.) produces the insulin

21. DA: mae fe'n mynd syth mewn i'r ardal pinc na (.) sef gwythïen (.) y gwaed
it goes straight into that pink area (.) which is the vein (.) the blood

22. DA: mae hwnna wedyn ny (.) yn gweithio ar y glwcos sy'n y gwaed (.) a

galluogi y glwcos i fynd mewn i'r celloedd
*that then (.) works on the glucose that's in the blood (.) and enables the
glucose to go into the cells he other nine*

23. DA: oni bai bod yr inswlin na (.) chi methu (.) defnyddio'r glwcos yn y gwaed
unless the insulin is there (.) you can't (.) use the glucose in the blood
24. DA: a mae pob cell yn y corff (.) fel ffatri
and every cell in the body (.) is like a factory
25. DA: i gael yr ynni (.) mae raid cael y petrol
to get the energy (.) they/you have to have the petrol
26. DA: a'r petrol yw'r glwcos
and the petrol is the glucose
27. DA: so os nag yw'r inswlin mae'r [?] glwcos yn y gwaed yn mynd lan
so if there isn't any [?] insulin the glucose in the blood goes up
28. DA: mae'n cael ei waredu (.) yn yr urine
it's expelled (.) in the urine
29. DA: chi'n (.) colli lot a lot o'r urine
you (.) lose a lot of the urine
30. DA: chi'n mynd yn flinedig tu hwnt tu hwnt
you become extremely extremely tired
31. DA: a oherwydd bod dim digon o glwcos i'r celloedd mae nhw'n crio allan i'r
corff (.) cynhyrchwch ragor (.) o glwcos
*and because there isn't enough glucose for the cells they cry out to the body
(.) "produce more (.) glucose"*
32. DA: felly mae'r afu yn trio cynhyrchu (.) glwcos
so the liver tries to produce (.) glucose
33. EJ: +< ie
yes
34. DA: a mae'r cyhyre (.) a'r braster yn trio cynhyrchu glwcos
and the muscles (.) and the fat try to produce glucose
35. DA: a chi'n colli pwyse

and you lose weight

36. DA: so y syched a'r blinder a'r colli pwyse (.) dyna [?] (.) dyna beth yw e
so the thirst and the tiredness and the weight loss (.) that's [?] (.) that's what it is
37. EJ: +< iawn (.) (guest name) (.) beth oedd eich profiad chi felly
alright (.) (guest name) (.) what was your experience then?
38. EJ: achos (.) math dau sy arnoch chi yn dyfe
because (.) you've got type two, haven't you
39. GU: +< ie [?]
yes
40. GU: wel fel oedd dact- doctor Ann yn sôn (.) blinder ofnadwy (.) oedd y
symptom gynta
well as doctor Ann was saying (.) terrible tiredness (.) was the first symptom
41. GU: um (.) a wedi by- bod yn mynd nôl a mlaen at y meddyg
um (.) and had been going back and forth to the doctor
42. GU: naeth neb (.) feddwl mai (.) clefyd y siwgr oedd arno i
nobody (.) thought that (.) I had diabetes
43. GU: o'n i'n dal i fod yn flinderus
I was still tired
44. GU: uh (.) mynd i'r ysgol
uh (.) going to school
45. GU: a fel athro cerdd (.) oedd gofynion arna i i weithio amser cinio
and as a music teacher (.) I was required to work during lunch time
46. GU: ond (.) o'n i'n cysgu amser cinio
but (.) I was asleep during lunch time
47. GU: a waeth waeth na ny (.) o'n i'n byta (.) siocled i gael trial (.) um cael nerth o
rywle yn dyfe
and worse than that (.) I was eating (.) chocolate to try to (.) um get energy from somewhere, isn't it
48. GU: ond oedd na'n hala fe'n waeth

but that was making it worse

49. EJ: wrth gwrs
of course
50. EJ: o o'ch chi'n sychedig iawn hefyd
were you very thirsty too?
51. GU: oh sychedig ofnadwy
oh terribly thirsty
52. GU: fydden i'n mynd i'r gwely (.) <a fydden i'n> [?] cysgu am ryw (.) wel (.) awr
(.) a wedyn ny xxx dihuno
*I'd go to bed (.) <and I'd> [?] sleep for about (.) well (.) an hour (.) and then
xxx wake up*
53. GU: a syched (.) wel (.) oedd fel bwyta dŵr chimod neu xx potel fawr o
lemonade
*and thirsty (.) well (.) it was like eating water, you know, or xx a large bottle
of lemonade*
54. EJ: <a oedd y> [?] lemonade ddim yn beth da chwaith
<and the> [?] lemonade wasn't [?] a good thing either
55. GU: wel wrth gwrs nag oedd
well of course, no
56. GU: ond o'n i ddim yn sylweddoli ny ar y pryd
but I didn't realise that at the time
57. EJ: felly am faint fuoch chi fel hyn cyn i nhw weud wel ie wy ni ni'n credu bod
(.) diabetes arnoch chi
*so for how long were you like this before they said "well yes, we think (.)
you've got diabetes"?*
58. GU: wel o'n i'n meddwl [?] bod fi di bod am flwyddyn
well I thought [?] I'd been a year
59. GU: yn diwedd wrth gwrs ffeindion nhw mas achos fe nes i lewygu
in the end of course, they found out because I fainted
60. EJ: reit
right

61. GU: ac wedyn ny roi prawf gwaed wedyn ny
and then gave a blood test then
62. EJ: a mae hwnna'n arferol
and that's normal?
63. EJ: dyna fel arfer mae pethe'n mynd
that's how things go usually?
64. DA: +< ody
yes
65. DA: mae mwyafrif o'r (.) clefydau ma ma t- math dau ydyn nhw wyth deg y cant
most of (.) these diseases, they're type two, eighty percent
66. DA: ryw ugain y cant math un
about twenty percent type one
67. DA: a mae math un yn taro plant fel arfer (.) neu yn [?] bobl yn ei arddege
and type one hits children usually (.) or people in their teens
68. DA: mae'n dod mlaen yn sydyn (.) fel na
it comes on suddenly (.) like that (clicks fingers of right hand)
69. DA: oherwydd mae'r pancreas yn methu yn gyfan gwbl
because the pancreas fails completely
70. DA: oherwydd mae na salwch autoimmune (.) le mae'r corff yn ymladd a (.) lladd (.) celloedd (.) y corff ei hunan
because there's an autoimmune illness (.) where the body fights and (.) kills (.) the body's own cells
71. DA: ond gyda hwn (.) mae hwn yn dod mlaen yn raddol
but with this (.) this comes on gradually (pointing to GU)
72. DA: a mae'r symptome'n gallu bod yn anelwig
and the symptoms can be ambiguous
73. DA: blinder (.) rywbeth niwlog iawn a gall fynd <ymlaen am flynydde> [?]
tiredness (.) something very hazy and can go <on for years> [?]

74. DA: yn y cyfamser of cwrs mae cymhlethdode (.) y diabetes (.) yn dechre
in the meantime of course, the complications (.) of the diabetes (.) begin
75. EJ: fi- t- dechreuoch chi (guest name) gyda (.) uh colli pwyse (.) wedyn ny
mynd ar ddiety ie
*you started, (guest name), with (.) uh losing weight (.) then went on a diet,
yes?*
76. GU: wel ne- nes i nore [=? ngore] i i i (.) i golli pwyse
well I did my best to (.) to lose weight
77. GU: ond oedd hwnnw ddim yn gweithio
but that wasn't working
78. EJ: ie
yes
79. GU: felly n- (.) es i ar dabledi wedyn ny
so (.) I went on tablets then
80. GU: o'n nhw ddim yn gweithio chwaith
they didn't work either
81. EJ: iawn
right
82. GU: es [?] i (.) es i ar inswlin
I went on insulin
83. EJ: ie
yes
84. GU: a (.) naeth pethe wella
and (.) things got better
85. GU: ac wedyn ny (.) o'n i'n sôn wrth doctor Ann fe ges i inswlin
and then (.) I was telling doctor Ann, I got insulin
86. GU: wy'n cymryd (.) inswlin sy'n (.) gweithio yn gyflym iawn (.) yn y dydd
I take (.) insulin that (.) works very fast (.) during the day
87. GU: y'ch chi'n cymryd e'n y bore amser cinio amser te
you take it in the morning, at lunch time and at tea time

88. GU: a chi'n bwyta bwyd (.) mewn pen hanner awr
and you eat food (.) within half an hour
89. GU: ac wedyn ny (.) y peth naeth wella i lot fawr (.) oedd yr inswlin sy dim ond
(.) yn cael ei ddydd- (.) ddefnyddio yn yr hwyr (.) y long-acting
*and then (.) the thing that made me a lot better (.) was the insulin that is
only (.) used in the evening (.) the long-acting (he reaches into the inside
pocket in his jacket and takes out a long white tube – a little bit bigger than
a pen)*
90. GU: a mae hwn wedi neud y tric i fi
and this has done the trick for me
91. EJ: a hwnna sy'n eich caniatáu chi siŵr o fod i gael cwsg (.) a ca- a peidio
goffod codi'n y nos a'r math na o beth i gyd
*and that's what allows you, probably, to sleep (.) and not to have to get up in
the night and all that kind of thing*
92. GU: +< wel ie (.) ie (.) xx yn bendant ie ie
well yes (.) yes (.) xx definitely, yes yes
93. EJ: felly [?] oes na ryw bethe falle sy yn (.) yn (.) wel yn y ne- yn y (.) yn
geneteg de felly
*so are there some things, perhaps that are (.) well in the (.) in the genetics
then?*
94. DA: mae na elfen genetic cryf iawn i'r math dau
there is a very strong genetic element to the type two
95. DA: gwedwch chi nawr bod gyda chi dou efaill a un yn cael (.) y math hyn o
diabetes
say now that you've got two twins and one gets (.) this type of diabetes
96. DA: mae'r efaill arall o dan risg o naw deg y cant o ddatblygu'r salwch
the other twin is at risk of ninety percent of developing the illness
97. DA: yn y math un (.) mae fe lot mwy gwan
in type one (.) it's a lot weaker
98. DA: dim ond ryw dri deg y cant (.) o'r ail efaill fydd yn cael y salwch
only about thirty percent (.) of the second twin will get the illness

99. DA: an- mae'r salwch (.) teip un yn tueddu digwydd mewn (.) clwsters
the illness (.) type one tends to happen in (.) clustures
100. DA: fydd lot mewn un ardal
there will be a lot in one area
101. DA: felly ni yn meddwl mai rywbeth yn yr amgylchedd (.) ynglwm â ryw (.)
ffactor gwan genetic
*so we think it's something in the environment (.) connected to some (.) weak
genetic factor*
102. DA: falle haint
perhaps an infection
103. DA: mae nhw di bod yn sôn am (.) virus coxsackie (.) rubella (.) neu glandular
fever
*they've been talking about (.) the coxsackie virus (.) rubella (.) or glandular
fever*
104. DA: bod rywbeth fel na (.) yn taro ar y pancreas ac
that something like that (.) hits the pancreas and...
105. DA: ond gyda (.) y yr ail fath ma (.) gordewdra (.) sy'n gysylltiedig fel arfer
gyda'r salwch yn enwedig (.) maint y wast
*but with (.) this second type (.) it's obesity [or perhaps being overweight] (.)
that's connected usually with the illness, especially (.) the size of the waist*
106. DA: tri deg pump (.) mwy na tri deg pump i ferched mwy na (.) pedwar deg i
ddynion
thirty-five (.) more than thirty-five for women, more than (.) forty for men
107. DA: na'r ffactor pwysig
that's the important factor
108. EJ: ryfedd
strange
109. DA: ie
yes
110. EJ: beth am y sgileffeithie nawr te gyda chi (guest name)
what about the side effects now then with you, (guest name)?

111. EJ: mae na sgileffeithie yn does
there are side effects, aren't there
112. GU: +< mae na sgileffeithie oes
there are side effects, yes
113. GU: wel wy'n diodde wrth (.) uh (.) beth mae nhw'n galw peripheral
neuropathy (.) uh colli ymwybyddiaeth yn y traed
*well I suffer from (.) uh (.) what they call peripheral neuropathy (.) uh loss of
consciousness in the feet*
114. GU: na un ffordd
that's one way
115. GU: neu (.) mae nhw'n llosgi'n annioddefol
or (.) they burn unbearably
116. EJ: pam mae [=? bod] hwnna'n digwydd felly
why does that happen then?
117. DA: be sy'n digwydd mae dau beth mae'r (.) mae'r (.) siwgr ma'n y gwaed (.) yn
achosi niwed (.) i'r arteries mawr a'r arteries bach bach
*what happens, there are two things the (.) this (.) sugar in the blood (.)
causes damage (.) to the big arteries and the little arteries*
118. DA: o ran yr arteries mawr chi o dan risg trawiad (.) stroc gwedwch (.) iawn
trawiad y galon neu stroc
*in terms of the big arteries, you're at risk of an attack (either heart attack or
stroke) (.) stroke, say (.) alright, a heart attack or a stroke*
119. DA: ond y rei bach ma sy'n neud y niwed anelwig
but it's these little ones that do the invisible [?] damage
120. DA: achos mae nhw'n gallu neud niwed i'r nerfe i gyd (.) yn enwedig nerfe'r
droed le chi'n colli (.) teimlad eich troed
*because they can damage all the nerves (.) especially the nerves of the feet
where you lose (.) feeling in your foot*
121. DA: le neu mae'r nerfe yn gorweithio
where...or the nerves overwork
122. DA: mae nhw'n llosgi'n ofnadwy
they burn terribly

123. DA: neu (.) mae'n gallu effeithio ar arennau
or (.) it can affect kidneys
124. DA: mae'n gallu effeithio ar lygad of cwrs (.) ac yr ymennydd
it can affect the eye, of course (.) and the brain
125. DA: fel o'ch chi'n gweud o'ch chi'n mynd yn anymwybodol
(addressing GU) as you were saying, you were losing consciousness
126. DA: felly chi'n goffod cael profion (.) cyson (.) blynnyddol (.) i weld ydy pwyse
gwaed yn mynd lan beth yw lefel y colesterol
*so you have to have (.) constant tests (.) every year (.) to see whether the
blood pressure is increasing, what's the level of cholesterol*
127. DA: felly chi'n goffod bod ar lluo o (.) um dabledi eraill i (.) gadw cymhlethdode
lawr
*so you have to be on a host of (.) um other tablets to (.) keep the
complications down*
128. EJ: +< xxx
129. EJ: mae mae mae'n anodd i chi (guest name) achos y'ch chi'n organydd yn
dy'ch
*it's very difficult for you, (guest name), because you're an organist, aren't
you*
130. DA: mmm mmm
131. EJ: chimod dach [?] chi'n goffod defnyddio'ch traed dipyn
you know, you have to use your feet quite a bit
132. GU: +< ody ody
yes, yes
133. GU: chi'n dod yn gyfarwydd â'r peth
you get used to it
134. GU: ond (.) wy'n (.) yn sylwi'n ddiweddar bod fi (.) ddim (.) cystal ag oedden i
yn dyfe
but (.) I've (.) realised recently that I'm (.) not (.) as good as I was, you know

135. GU: <y fi> [?] mynd yn hŷn
I'm [?] getting older
136. GU: a (.) os oes dim teimlad yn y traed wel feth- fe allwch chi ddim ware'r pedale
and (.) if there's no feeling in the feet, well, you can't play the pedals
137. EJ: chi'n cael probleme cerdded o'r herwydd
do you have problems walking because of it?
138. GU: dim probleme cerdded ond (.) os dw dw i uh yn colli teimlad (.) be sa glass neu rywbeth neu ryw wydr (.) chi'n gwybod yn ar y ffordd
not problems walking but (.) if I uh lose feeling (.) what if there was glass or something, or glass (.) you know on the road?
139. GU: fydden i ddim yn gallu teimlo bod fi di cerdded ar fe
I wouldn't be able to feel that I'd walked on it
140. GU: na be sy'n pryderu fi chi'n gwybod
that's what worries me, you know
141. DA: na pam mae reolaeth y diabetes yn goffod bod yn dynn iawn
that's why control of diabetes has to be very tight
142. DA: a chi fod o dan gofal nyrs arbennig a'r meddyg a cael profion unwaith neu dwywaith y flwyddyn ar (.) y traed pwyse gwaed y llyged yn gyson
and you're supposed to be under the care of a special nurse and the doctor, and have tests once or twice a year on (.) the feet, blood pressure, the eyes, regularly
143. DA: a mesur siwgr eich gwaed eich hunan
and measure your own blood sugar
144. DA: a mae hwnna'n niwsans yn dyw e
and that's a nuisance, isn't it
145. GU: mmm ody wel
mmm, yes, well...
146. DA: a chi'n goffod neud e unwaith neu dwywaith y diwrnod cyn pryd o fwyd (.) i weld beth yw'r siwgr i gael y (.) yr inswlin ma hefyd y'ch chi'n gael yn dy'ch chi

and you have to do it once or twice a day, before a meal (.) to see what the sugar is to get this (.) this insulin too that you're getting, aren't you

147. GU: +< mmm (.) mmm (.) ie
mmm (.) mmm (.) yes

148. GU: wy'n neud e unwaith
I do it once

149. EJ: ie
yes

150. GU: ond mae mae (.) mae'r ofal y gofal y'n ni'n gael yn yr ysbyty lleol fan hyn (.)
yn wych
but the (.) the care...the care we get in the local hospital here (.) is excellent

151. DA: +< <ond mae e'n niwsans> [?]
<but it is a nuisance> [?]

152. EJ: xxx falch o glywed hynny
xxx glad to hear that

153. GU: mmm mmm

154. EJ: wel diolch yn fawr iawn i chi
well thank you very much

155. EJ: a phob hwyl i chi (guest name)
and all the best to you, (guest name)

156. GU: +< croeso i chi
you're welcome

157. GU: diolch
thank you

158. EJ: diolch yn fawr iawn a chithe Ann hefyd diolch
thank you very much and to you too, Ann, thanks

159. DA: +< mmm

APPENDIX THREE:

Sample transcript – Housing Association data

Transcription conventions here are different to the ones I've used for the presentation of data excerpts in the body of the thesis. Specifically, here transcription of overlap and sequence is more rough than in the body of the thesis, where the presentation of sequence is potentially more important for analysis. Here overlap is marked by +< at the beginning of a line, which marks that the following text overlaps with part or all of the preceding line. Because lines here are usually full turns (and hence sometimes very long), there can be a series of overlaps to a preceding line.

Voids

Sample from: *Voids* recording, lines 79-148 00:08:16 – 00:13:09
Participants: Pam, Nell, Bronwyn, Danni, Jim, Phoebe, Cheryl, Paul, Beth, Beverly
Situation: Meeting of the *Gweithgor Foidiau* ('Voids Working Group') around table in committee room
Note: SXX marks a voice that I can't place
Note: They discuss a lot of properties in this recording. I've transcribed any numbers (flat numbers and house numbers) but obscured any house names, building names, street names or place names. I have also obscured local shop names, but not commercial names that are well-known across the country

79. Nell: ia (.) ond um (.) so (.) a dyn (place name) a'r (property name)
wedyn (.) timod yr (property name) a dyn (property name) xx
(place name) neu (place name)
yes (.) but um (.) so (.) and then (place name) and the (property name) then (.) you know the (property name), and then (property name) xx (place name) or (place name)?
80. SXX: (place name)
81. Nell: ti'n (.) xxx
you (.) xxx
82. Jim: ma hwnnw wag eto dy
that's vacant again, isn't it
83. Nell: be di dod yn wag yn ôl eto

what, it's become vacant again?

84. Jim: xxx

85. Pam: fanno gaeth raid
is that the place that got raided?

86. Jim: ia
yes

87. SXX: xxx

88. Phoebe: +< fanna ma um
is that where, um..?

89. Jim: ia (.) pa run di [?] (.) hwnnw xx (.) xxx wan ia
yes (.) which one is (.) that xx (.) xxx now, yes?

90. Pam: ie

91. Jim: <ie ond> [?] dydy hwnnw'm yn wag
yes but that's not vacant

92. Phoebe: nadi mae dal yna mae di cael raid de
no, it's still there, it got raided, right

93. Jim: yndy oh ie ma hwnnw di oso- ma hwnnw di osod do
yes, oh yes, that's been let, hasn't it

94. Phoebe: do do ma dyna e- eiddo di cael ei <gosod di hein i gyd> [?]
yes, that's... these are all properties that have been let

95. SXX: +< xxx

96. Jim: oh (.) s-

97. Nell: ma heina sy dal yn wag dal yn ticio dydy
those that are still vacant are still ticking, aren't they

98. (laughter)

99. Beth: jyst xxx
just xxx

100. Nell: +< <ma rheina sy di> [?] cael ei gosod xxx
those that have been let xxx
101. Phoebe: +< xxx
102. Jim: +< xxx di cael ei osod ond um (.) xxx
xxx has been let but um (.) xxx
103. Phoebe: +< do
yes
104. Beth: +< xxx hir ie
xxx long, yes
105. Phoebe: be ddigwyddodd gaethon ni um enwebiada gyn y cyngor (.) a pobol
lleol up in arms oeddau nhw di clywad xxx achlust bod y bobol (.) a-
um annymunol ma di cael cynnig tenantiaeth (.) um (.) ond ond
oedd gyn y cyngor neb arall (.) dyn mi ddaru ni gynnig y
denantiaeth (.) iddyn nhw'n diwadd (.) <a wedyn ma nhw di cael
raid wan> [?] (.) (laughs)
*what happened, we got um nominations from the council (.) and the
local people up in arms, they'd heard xxx a rumour that these (.)
unpleasant people had been offered a tenancy (.) um (.) but the
council didn't have anybody else (.) so we offered them the tenancy (.)
in the end (.) and then they've been raided now (.) (laughs)*
106. Jim: +< do
yes
107. Beth: +< so pobol lleol yn iawn
so the local people were right
108. Phoebe: oeddau
yes
109. Jim: xxx
110. Nell: ond (.) amlwg ma nhw angan tai
but (.) obviously they need houses
111. Beth: mmm

112. Nell: um (.) wedyn (property name) eto an- anodd i osod oedd (property name) ynde (.) eiddo eto (.) ma hwnna rwbath dan ni isho edrych arno fo (.) a wedyn (.) pump (property name) oedd na waith (.) gwaith oedd yn (property name)
um (.) so (property name) again, (property name) was difficult to let, wasn't it (.) property again (.) that's something we want to look at (.) and then (.) five (property name), there was work (.) was it work in (property name)?
113. SXX: na
114. Beth: nag oedd oedd o'n jyst anodd i osod dw meddwl xxx
no, it was just difficult to let, I think xxx
115. Phoebe: oedd [?] gaethon ni xxx gwrthod
yes, we got xxx rejected
116. Beth: do
yes
117. Phoebe: oedd na un (.) yn gwrthod
there was one (.) who refused...
118. Beth: +< y [?] grisia
the stairs
119. Phoebe: +< xxx (.) grisia ia
xxx (.) stairs yes
120. Beth: ia ia
121. Nell: so dy- dyna oedd yr (.) eiddo sy di cael ei gosod hyd at wan sy di cymryd drost wyth-deg o ddyddia
so those were the (.) properties that have been let up to now that have taken over eighty days
122. Beth: mmm
123. Nell: pan awn ni fewn i eiddo anodd i osod (.) ma na gynnon ni amryw o rhei drost wyth-deg yn wag gynnon ni so nawn ni bigo hwnna i fyny (.) um (.) wan (.) oedd oedd (.) dan ni di bod yn cael snapshot really (.) um er mwyn cael gweld os oedd na (.) um (.) efo gwahanol ardaloedd wrth gwrs hwyrach ma hwnna'n rwbath dan ni isho

edrych arno fo (.) os dan ni isho edrych efo'r de a'r gogledd
 hwyrach (.) ond (.) <oeddau ni> [=? oeddwn i] hefyd isho gweld os
 oedd na issues <ynglŷn â> [?] enwebiada (.) os oedd hwnna'n codi
 (.) yn hwnna ond (.) mae'r system Lean (.) yn mynd i bigo lot o
 hynny fyny i chi beth bynnag dydy (.) so um (.) ma na (.) ma na dal
 issue (.) efo (.) um (.) eiddo newydd (.) lle os di rywun di bwydo i
 fewn cyn iddo fo fod yn barod (.) bod y lle mynd i fod yn (.) lettable
 (.) a dydan ni'm yn gosod o tan wedyn (.) ia
*when we go to property that's difficult to let (.) we've got a number of
 ones over eighty days that are vacant, so we'll pick that up (.) um (.)
 now (.) we've been getting a snapshot really (.) um in order to see if
 there was (.) um (.) with different areas, of course, perhaps there's
 something we want to look at (.) if we want to look with the south
 and north perhaps (.) but we also wanted to see if there were issues
 about nominations (.) if that arose (.) in that, but (.) the Lean system
 (.) is going to pick up a lot of that for you anyway, isn't it (.) so um (.)
 there is (.) there is still an issue (.) with (.) um (.) new property (.)
 where if somebody has inputted before it's ready (.) that the place is
 going to be (.) lettable (.) and that we don't let it til later (.) yes*

124. Beth: mis Ionawr
 January

125. Nell: ynde (.) dyn nhw'm yn gallu unwaith ma'r dyddiad na deud fel
 enghraifft wan bod oedd (.) y tai yn (.) yn (property name) reit di
 cael ei deud bod o'n barod mis Rhagfyr (.) a nathon nii ddim gosod
 nhw tan mis Ionawr (.) unwaith oedd o mynd drost y diwrnod
 Rhagfyr yna (.) oedd o'n dechra tician (.) ma rhaid i criw (.)
 datblygu (.) tynnu'r tic na off on- o- oedd o di cael ei neud ond ma'n
 amlwg (.) dw'm yn deud ydy
*right (.) they can't...once that date...say as an example now that (.) the
 houses (.) in (property name), right, had been said that it's ready in
 December (.) and we didn't let them until January (.) once it went
 over that December day (.) it started to tick (.) the development (.)
 team (.) tave to take that tick off, it had been done, but it's obvious (.)
 I'm not saying that...*

126. Phoebe: ma'n hawdd (.) neud o
 it's easy (.) to do

127. Nell: ti sôn am gant dau-ddeg (.) um (.) diwrnod jyst ar gyfar tri eiddo
 fanna (.) sy'n ychwanegu (.) ar nifer o ddyddia dydy (.) um timod (.)
 ma rhaid ni neud siwr so ma angan (.) unwaith eto i i (.) i atgoffa (.)

datblygu o'r drefn sa ti'n c- (.) rhoid uh enw fi'n erbyn [?] hwnna (.)
 plis (.) ti'n gwobod um (.) mae o yn effeithio dydy dy- dy- dy- dy- wel
 dio'm y gwiriona nadi
*you're talking about a hundred and twenty (.) um (.) days just for
 three properties there (.) which adds (.) to the number of days, doesn't
 it (.) um you know (.) we have to make sure so we need (.) again to (.)
 to remind (.) development of the procedure, could you (.) put um my
 name against that (.) please (.) you know um (.) it affects, doesn't it,
 well it's not the most truthful [?], is it*

128. Beth: dio ddim yn wir
it's not true

129. Cheryl: so be ma nhw'n neud bob tro ma nhw meddwl bod na dai'n dod
 drosodd [?] ia
*so what do they do, every time they think there are houses coming
 over, yes?*

130. Nell: pan ma n- (.) pan ma nhw'n bwydo'r gwybodaeth i fewn ma nhw'n
 gorfod roid i fewn pryd mae o mynd i fod yn barod (.) ac yn aml
 iawn mae'r cytundeb yn dod i ben ond efo (property name) (.) gan
 bod oedd y grant oeddann ni'n dynnu lawr ar i gyfar o oeddann ni'm
 yn gallu tynnu fo lawr tan dechra flwyddyn (.) nath (company
 name) (.) gytuno (.) i gadw fo tan ddechra flwyddyn do (.) a dyn
 wrth gwrs ma di cyfri hwnna dydy (.) er bod y gontract di dod i ben
 (.) oedd na gytundab so (.) ond oeddann ni di penderfynu bod doedd
 oeddann ni'm yn mynd i oeddann ni'm yn gallu osod o oeddann ni'm yn
 gallu tynnu grant i lawr ti gweld (.) tan dechra flwyddyn (.) felly
 dyna sut ma di [?] gweithio (.) a ma'n jyst dangos dydy ma'r
 trafodaethau eraill yn amharu
*when (.) when they input the information, they have to input when it's
 going to be ready (.) and very often the agreement comes to an end,
 but with (property name) (.) because there was the grant we were
 taking down for it, we couldn't take it down until the start of the year
 (.) (company name) (.) agreed (.) to keep it until the beginning of the
 year, didn't they (.) and then of course it's counted that, hasn't it (.)
 although the contract has come to an end (.) there was an agreement,
 so (.) but we'd decided that we weren't going to...we couldn't let it, we
 couldn't take down the grant, you see (.) until the beginning of the
 year (.) so that's how it's worked (.) and it just shows, the other
 discussions interfere*

131. Beth: ar y

with the...

132. Nell: ar stwff ni de=
with our stuff, you see
133. Phoebe: = <ond ma hwn y> [?] tri diwrnod bob un (.) yn (property name) (.)
os di'r diwrnod cwblhad ymarferol ar y (.) dydd Gwenar wel da- (.)
dan ni'n cael (.) foidio ond dan ni'm yn (.) dan ni'm gallu osod o tan
y dydd Llun (.) o ran mesur perfformiad dio'm yn dangos xxx
*<but this is the> [?]three days each (.) at (property name) (.) if the
practical [?] completion day is on the (.) Friday, well (.) we can (.)
void, but we can't (.) we can't let it until the Monday (.) in terms of
measuring performance it doesn't show xxx*
134. Beth: dim guarantee os y bora dydd Llun chwaith (.) os dan ni'm yn cael o
tan pump (.) does a neb yma i bigo fo fyny (.) tan dan ni dod i fewn
dydd Llun (.) a wedyn ffonio'r pobol sydd fod i gael y lle lle xxx
*no guarantee if the Monday morning either (.) if we don't get it until
five (.) there's nobody here to pick it up (.) until we come in on
Monday (.) and then phone the people who are supposed to get the
place, where xxx*
135. Phoebe: +< cael gafael arnyn nhw> [?] ia
get hold of them, yes
136. Beth: xxx
137. Phoebe: +< tai newydd di hein i gyd
these are all new houses
138. SXX: +< xxx
139. Beth: ah dan ni'n dal ar tai newydd
ah, we're still on new houses
140. Phoebe: efo tai newydd yn (property name)
with new houses at (property name)
141. Beth: ia
142. Phoebe: ond ma nhw xx (.) oeddan ni'm yn gallu gosod o tan dydd Llun nag
oeddan
but they're xx (.) we couldn't let it until Monday, could we

143. Beth: nacdan (.) cofio
no (.) remember
144. Phoebe: +< ma'r (.) dyddiad (.) xxx (.) bod nhw'n foid (.) di mynd tri diwrnod
cynt sy sy ddim yn (.) adlewyrchu'r perfformiad nadi
*the (.) date (.) xxx (.) that they're void (.) has passed three days
previously, which doesn't (.) reflect the performance, does it*
145. Beth: +< do
yes
146. (.)
147. Nell: so unwaith eto hwyrach fydd raid ni edrych ar neud manual (.)
adjustment ar gyfar hynny (.) um Danni
*so once again, perhaps we'll have to look at doing a manual (.)
adjustment for that (.) um Danni*
148. Beth: +< mmm

End of sample transcript

APPENDIX FOUR:

Terms in *Syrjeri* data according to codification category**Codified:**

Term type	English	Tokens
afu	liver	12
ambiwylans	ambulance	5
anymwybodol	unconscious	1
aren	kidney	8
bacteria	bacteria	1
beichiogrwydd	pregnancy	1
bol	belly/stomach	1
braich	arm	4
braster	fat	3
brech	rash	7
brechu	vaccinate	2
bys	finger	2
calon	heart	3
camera	camera	3
carthion	excrement	2
cefn	back	1
cell	cell	7
cemotherapi	chemotherapy	1
claf	patient	1
clefyd	disease	1
coes	leg	4
colesterol	cholesterol	1
corff	body	19
croen	skin	5
cyhwr	muscle	1
cylchrediad	circulation	1
cymal	joint	2
diagnosis	diagnosis	3
geneteg	genetics	1
glwcos	glucose	11
gordewdra	obesity (lit. state of being too fat)	1
gwaed	blood	11
gwddf	neck	1
gwres	temperature, heat	2
haint	infection	1
heintus	infectious	2
hunanimiwn	autoimmune	1

imiwnedd	immunity	1
inswlin	insulin	11
llaw	hand	4
llawfeddyg	surgeon	3
lleddfu	alleviate	1
llid yr ymennydd	meningitis	6
llygad	eye	2
meddyg	doctor	8
moddion	medicine	1
nerf	nerve	3
nyrs	nurse	1
organ	organ	4
poen	pain	9
poen yn y bol	stomach ache	2
prawf	test	7
prawf gwaed	blood test	2
salwch	illness	9
stent	stent	3
steroid	steroid	4
stroc	stroke	2
stumog	stomach	1
symptom	symptom	15
tabled	tablet	6
tost	ill, sick	2
tostrwydd	illness	1
trawiad	attack, stroke	1
trin	treat	1
triniaeth	treatment	10
troed	foot	9
tymheredd	temperature	2
uwchsain	ultrasound	1
ymennydd	brain	2
ymwybyddiaeth	consciousness	1
ysbyty	hospital	12
ysgwydd	shoulder	1
ysgyfaint	lungs	1

Total codified terms: 73

Total codified tokens: 282

Not codified:

Term type	English	Tokens
autoimmune	autoimmune	2
bile duct	bile duct	1
bogell	belly button/navel	1

briws	bruise? Sore?	2
consciousness	consciousness	1
consultant	consultant	3
erythema	erythema	1
gallbladder	gallbladder	1
gallstone	gallstone	1
glandular fever	glandular fever	1
intensive care	intensive care	1
jaundice	jaundice	1
joint	joint	1
keyhole	keyhole	1
kidney	kidney	3
rash	rash	5
rehab	rehabilitation	1
rheumatoid arthritis	rheumatoid arthritis	1
rheumatologist	rheumatologist	1
salwch y blaidd	lupus	1
sic	sick	2
SLE	SLE	3
straen	strain (of bacteria)	3
suppresso	suppress	1
taflu nôl	throw up	1
ultrasound	ultrasound	1

Total not codified terms: 26

Total not codified tokens: 41

Partly codified:

Term type	English	Tokens
adran damweinie	accidents (and emergencies) department	1
aflwydd	illness?	1
brechiad	vaccine/vaccination	4
bustl	gallbladder	8
damweinie	accidents (and emergencies department)	2
genetic	genetic	2
gwrthbiotigau	antibiotics	1
gwyneb	face	1
gwythien	vein	6
Immunoglobulin	immunoglobulin	1
meningococcus	meningococcus	2
paracetamol	paracetamol	2
pwyse'r gwaed	blood pressure	1
rubella	rubella	1

trawiad y galon	heart attack	1
virus	virus	1

Total partly codified terms: 16 Total partly codified tokens: 35

Mixed:

Term type	English	Tokens
arteri	artery, blood vessel, (poss. vein?)	4
comesgor	miscarry	1
cerrig y bustl	gallstones	1
cholecystitis	cholecystitis	1
clefyd melys	diabetes	1
clefyd y siwgr	diabetes	2
cyflwr	condition	6
diabetes	diabetes	5
doctor	doctor	9
dos	dose	1
gwrthgorff	antibody	6
llawdriniaeth	operation	5
lupus	lupus	6
meningitis	meningitis	6
pancreas	pancreas	6
pen tost	headache	1
perfedd	intestine/bowel	3
pwyse gwaed	blood pressure	2
sâl	ill, sick	1
syrjery	surgery (doctor's, dentist's, etc.)	3
systemic	systemic	1
treto	treat	1
urine	urine	2
wast	waist	1

Total mixed terms: 24 Total mixed tokens: 75

Not applicable:

Term type	English	Tokens
ail fath	second type of diabetes	1
butterfly rash	butterfly rash	1
chest infection	chest infection	1
discoïd lupus	discoïd lupus	1
dwythen fach	small duct	1
dwythen fawr	large duct	4

fflw moch	swine flu	2
haemophilus	haemophilus	1
keyhole surgery	keyhole surgery	2
llawdriniaeth camera	keyhole surgery	2
long-acting	long-acting	1
math dau	type two	5
math un	type one	5
meningitis #by	meningitis B	1
meningococcus #by	meningococcus B	1
meningococcus #cy	meningococcus C	3
peripheral neuropathy	peripheral neuropathy	1
Raynauds	Raynauds	1
systemic lupus erythematosus	systemic lupus erythematosus	1
teip un	type one	1
triniaeth keyhole	keyhole surgery	1
virus coxsackie	coxsackie virus	1

Total not applicable terms: 22 Total not applicable tokens: 38

Total all terms: 161 Total all tokens: 471

APPENDIX FIVE:

Terms in Housing Association data according to codification category

Codified:

Term type	English	Tokens
addasu	adapt / modify	2
adrodd	report (v)	3
adrodd yn ôl	report back	2
adroddiad	report (n)	3
amserlen	timetable, schedule	2
archeb	order (n)	6
archebu	order (v)	1
argyfwng	emergency	2
asesiad	assessment	1
bin	bin	2
blaenoriaeth	priority	1
bloc	block (of flats)	2
cais	application	1
cegin	kitchen	31
clerc	clerk	2
cofrestr	register (n)	1
cost	cost	2
cwblhau	complete (v)	1
cwmni	company (business)	10
cydymffurfio	comply, conform	2
cyfanswm	total	1
cyfarfod (n)	meeting	8
cyfarfod (v)	meet	3
cyfathrebu	communicate	1
cyflenwi	supply (v)	1
cyflenwr	supplier	1
cyfranogiad	participation	1
cyfrifol	responsible	2
cyfrifoldeb	responsibility	4
cyhoeddiad	publication	1
cyllidol	fiscal / financial	1
cymradwyo	approve	1
Cyngor	Council	7
Cyngor Sir	County Council	2
cynllun (drawing)	plan, design (n)	4
cynllun (project)	plan, scheme (n)	3
derbynfa	reception	1

desg	desk	1
digartref	homeless	1
diogelwch	safety, security	4
dod i ben (intr)	expire	1
dodrefn	furniture	1
drafft	draft (first copy)	2
dror	drawer	1
drws	door	4
drws nesaf	next door	2
dylunio	design (v)	1
enwebiad	nomination	6
ffenestr	window	9
fflat	flat (apartment)	8
ffonio	telephone (v)	1
fformat	format (n)	2
ffurflen	form (n)	4
ffurflen gais	application form	1
gardd	garden	1
gorffeniad	finish (n)	1
gosod	let (v)	10
gosodiad	layout	1
gwag	empty / vacant	1
gwasanaeth	service	1
gwedd	phase	3
gweithlu	workforce	1
heddlu	police	2
iechyd a diogelwch	health and safety	1
lansio	launch (v)	2
llawlyfr	handbook, manual	4
llawr	floor	1
llythyr	letter	4
panel (group)	panel (group)	1
panel (sheet)	panel (sheet)	1
penderfyniad	decision	1
penodi	appoint	3
pensaer	architect	2
personel	personnel	1
plismon	policeman	2
polisi	policy	4
pris	price	2
rhestr	list	9
rhestr aros	waiting list	2
rhyddhau	release / discharge (v)	5
risg	risk (n)	3
safle	site	2
safon	standard	1
stad	estate	14

stafell	room	7
stafell wely	bedroom	2
staff	staff	4
swyddfa	office	3
swyddog	officer	2
swyddogol	official	1
sylwadau	comments	2
system	system	1
taflen	leaflet	1
tâl	payment	1
talw	pay (v)	6
targed	target	2
tenant	tenant	9
tim	team	2
tir	land	4
trafod	discuss	4
trefnu	arrange	3
trwsio	repair (v)	2
tŷ	house	12
uned	unit	1
ymgyngori	consult	1
ymgyngoriad	consultation	1
ymweliad	visit	5
yswariant	insurance	1
Total codified terms: 108		Total codified tokens: 331

Not codified:

Term type	English	Tokens
access	access (n)	1
ailddylunio	redesign	1
beading	beading	1
bildio	build (v)	1
building (insurance)	building (insurance)	1
checio	check (verify)	3
check	check (verification)	2
clerc gwaith	clerk of works	2
comment	comment	1
compound	compound (designated area)	1
contents	contents	2
date of birth	date of birth	1
decoratio	decorate	1
decoration	decoration	1
decorator	decorator	1

due	due	1
email	email	1
glass	glass	1
guideline	guideline	1
health and safety	health and safety	2
home visit	home visit	1
incident	incident	1
internal decoration	internal decoration	1
issue	issue	15
kitchen	kitchen	1
knock-on	knock-on	1
latch	latch	1
mock-up	mock-up	2
open plan	open plan	1
police	police	1
pressure	pressure	1
reference	reference	1
reference number	reference number	1
rhoi ymlaen	put forward / propose	3
seinio off	sign off	1
setio i fyny	set up (v)	2
shortlistio	shortlist (v)	1
sortio	sort (v)	6
sortio allan	sort out	2
specification	specification	1
storage	storage	4
supplyio	supply (v)	1
suspend	suspend	3
symud i mewn	move in	2
tâl gwasanaeth	service charge	1

Total not codified terms: 45

Total not codified tokens: 81

Partly codified:

Term type	English	Tokens
archwilydd	investigator, auditor, inspector	1
argraffwr	printer	2
cytundeb	contract, agreement	1
gwydr	glass	5
llawlyfr	handbook, manual	2

Total partly codified terms: 5

Total partly codified tokens: 11

Mixed:

Term type	English	Tokens
asesiad risg	risk assessment	1
bathroom	bathroom	1
cooker	cooker	1
cyfranogi	participate	1
cynnal a chadw	maintenance	6
dod i ben (tr)	annul	1
dyluniad	design (n)	7
insiwrans	insurance	2
job (position)	job (position)	3
job (task)	job (task)	4
llofft	bedroom	4
opsiwn	option	4
plan	plan (n)	1
pres	money	1
printio	print (v)	2
strwythur	structure	1
treial	trial (n)	2
tynnu nôl	pull out / withdraw	1

Total mixed terms: 18**Total mixed tokens: 43****Not applicable:**

Term type	English	Tokens
cofrestr ar y cyd	joint register	1
contents (insurance)	contents (insurance)	2
flyer	flyer	2
on board	on board	1

Total not applicable terms: 4**Total not applicable tokens: 6****Total all terms: 180****Total all tokens: 472**

