Complementary perspectives on hedging behaviour in family discourse: The analytical synergy of variational pragmatics and corpus linguistics

Brian Clancy

Mary Immaculate College, University of Limerick, Ireland

This paper argues that corpus linguistics offers a methodology which benefits variational pragmatic analysis in a number of ways. Corpus linguistic tools such as word frequency lists allow the researcher to construct a detailed 'pragmatic profile' of a word, cluster or act. This, coupled with the fact that most corpora are constructed to be representative of a particular language variety, facilitates an accurate account of language-use differences across various social categories. Pragmatic analysis relies heavily on context for its interpretation. Therefore, an illustrative case study of two corpora representing spoken language recorded in the home environment, one from a middle class Irish family and one from a family from the Irish Traveller Community will be utilised in order to elucidate the benefits of the synergy of corpus linguistics and variational pragmatics. Specifically, the variational distribution of the occurrences of hedges across these two distinct cultural groupings will be examined.

1.0 Introduction

Variational pragmatics was first proposed as an analytical framework by Schneider and Barron (2005) in order to address research gaps that existed in both modern dialectology and pragmatics. According to Schneider and Barron (2008: 1), variational pragmatics 'investigates pragmatic variation in (geographical and social) space.' Furthermore, Barron and Schneider (2009: 426) maintain that variational pragmatics 'investigates intra-lingual differences i.e. pragmatic variation between and across L1 varieties of the same language.' It is concerned with how the choice of one pragmatic strategy over another encodes macro-social indices of region, socio-economic status, ethnicity, gender or age in everyday language use. However, this is not to suggest that these five types are a closed set; the impact of other macro-social factors such as education and religion can also form part of this research framework. In addition, various micro-social factors, for example, power and social distance or register which impact on language variation can also be considered. However, in terms of a practical research agenda, Schneider and Barron (2008: 18) suggest that:

> Currently, variational pragmatics concentrates primarily on macro-social variation. It aims at determining the influence of each macro-social factor on language use individually…At a later stage it will be necessary to systematically include micro-social variation and to investigate the interaction between micro-social and macro-social factors.

The impact of both macro- and micro-social factors on pragmatic choice is essential to our understanding of language-use differences. Wolfram and Schilling-Estes (2006: 93) state that

'knowledge of when and how to use certain forms is just as important for communication as the literal understanding of structures and words.' However, they acknowledge that the study of how language is used in context is a relatively recent development in dialectology, especially when compared to the traditional focus on language form (pronunciation, vocabulary and grammar). In terms of the study of pragmatics, two criticisms of contemporary, cross-cultural pragmatics are posited by Schneider and Barron (2008), mainly in relation to the degree of representativeness of these studies. The first is that these studies are based on the assumption that language communities of native speakers are homogenous wholes when language variation is considered, thus, in a sense, negating the impact of social variables on language communities. In addition to this, Schneider and Barron (*ibid*.) claim that many researchers in this area employ participants from student communities, often from their own courses, thereby further compromising representativeness. While these studies are undoubtedly insightful, this lack of representativeness makes it difficult to formulate reliable generalisations about typical language use.

Hence, in general, Schneider and Barron maintain that studies into pragmatic variation can be criticised in relation to both their scope and representativeness; however, as exceptions they cite two studies that concentrate on regional language variation in English. These studies, Tottie (1991) and McCarthy (2002), are corpus-based studies. Both of these studies focus on the differences between backchannels (or response tokens) in British and American English. Tottie employs the London Lund Corpus (LLC) and the Santa Barbara Corpus (CSAE), and McCarthy the Cambridge and Nottingham Corpus of Discourse in English (CANCODE) in addition to a similar-sized sample of the Cambridge North American Spoken Corpus (CNASC). McCarthy (2002) maintains that cross-corpora comparisons of different varieties of the same language are

useful for a number of reasons. Crucially for the study of variational pragmatics, he notes that they provide safer ground for generalisations – all four corpora employed by Tottie and McCarthy have been specifically designed to represent standard British (LLC and CANCODE) and American (CSAE and CNASC) English, thereby alleviating some of the criticisms aimed at cross-cultural pragmatic research. Indeed, one of the strengths of corpus linguistics is that it has long been concerned with issues of representativeness; and, while the issue has never been resolved perhaps, this has resulted in an approximate but fairly robust approach to how 'representative' might be construed in the realm of language study (see Atkins *et al.* 1992, Clear 1992, Biber 1993, Crowdy 1994, Tognini-Bonelli 2001, Hunston 2002, McEnery *et al.* 2006).

1.1 The synergy of variational pragmatics and corpus linguistic methodology

Corpus linguistics offers a methodology which benefits the study of variational pragmatics in a number of ways. Jautz (2008: 146) maintains that another benefit that corpora offer the variational pragmatist is 'large amounts of naturally-occurring data, i.e. language in use, but also large amounts of comparable data from different varieties of one language.' In addressing this benefit, in relation to the synergy of critical discourse analysis (CDA) and corpus linguistic methodology, Orpin (2005: 39) cautions that 'an attendant danger in using a large corpus is that the researcher may feel swamped by the huge amount of data s/he is faced with.' She maintains that a good entry point for researchers in CDA is the corpus frequency list and this also holds true for variational pragmatics. When applied to the study of variational pragmatics, a word frequency list may allow the identification of items that may be characteristic of the pragmatic system of a particular language variety. For example, Table 1 shows a direct comparison of the

frequency lists for the top 25 words of the spoken component of the British National Corpus (BNC) and the Limerick Corpus of Irish English (LCIE)[1]:

Table 1: Top 25 word frequency counts for the BNC and LCIE[2]

	BNC	LCIE
1	*the*	*the*
2	*I*	*I*
3	*you*	*and*
4	*and*	*you*
5	*it*	*to*
6	*that*	*it*
7	*a*	*a*
8	*'s*	*that*
9	*to*	*of*
10	*of*	*yeah*
11	*n't*	*in*
12	*in*	*was*
13	*we*	*is*
14	*is*	*like*
15	*do*	*know*
16	*they*	*he*
17	*er*	*on*
18	*was*	*they*
19	*yeah*	*have*
20	*have*	*there*

21	*what*	*no*
22	*he*	*but*
23	*to*	*for*
24	*but*	*be*

Table 1, in addition to highlighting some *potential* pragmatic similarities between British and Irish English, also may point toward likely differences. For example, the personal pronouns *I* and *you* (positions two and three in the BNC and two and four in LCIE), characteristic of the deictic system in many languages, feature prominently and appear to be comparable (cf. Plevoets *et al.*'s 2008 study of pronouns in Netherlandic and Belgian Dutch, for example). In contrast, the pronoun *we* is in 13th position in the BNC but does not appear in the top 25 words in LCIE. Furthermore, the response token *yeah* occurs in 10th position in LCIE but in 19th in the BNC. This could indicate a predominance of *yeah* as a response token in informal, spoken Irish English as opposed to British English as represented in the spoken component of the BNC (see also O'Keeffe and Adolphs, 2008). Finally, two tokens with the potential to hedge in Irish English, *like* and *know*, both appear in LCIE (positions 14 and 15 respectively) but do not feature within the top 25 items on the BNC list. This may indicate that speakers of Irish English hedge more than their British counterparts or that these hedges take different forms. Both the hedges *like* and *you know* are amongst those explored further in the case study presented here.

Corpus word frequency lists are, admittedly, a raw measure of comparability, based on, as Table 1 demonstrates, the potential of a word form rather than its actual function. Jautz (2008) examines the BNC and the Wellington Corpus of Spoken New Zealand English for expressions of gratitude in British and New Zealand radio phone-in and broadcast interviews. She comments

that based solely on frequency, there are 287 expressions of gratitude in the British corpus and 129 in the New Zealand corpus, suggesting that the British are more polite because they use more expressions of gratitude. However, Jautz demonstrates that when these expressions are analysed more closely, the opposite appears to be the case. She found that New Zealanders take more care in phrasing their expressions of gratitude and name a reason for their gratitude more often than the British. Therefore, she argues, New Zealanders attend more to the face wants of their addressees offering 'a small piece of evidence that New Zealanders are in fact more polite than the British in this respect' (p. 170). Similarly, Farr and O'Keeffe (2002) examine the occurrences of the hedges *I would say* and *I'd say* in three spoken corpora: LCIE, CANCODE and a corpus of American spoken data from the Cambridge International Corpus. They found that these hedges are used more frequently by Irish speakers than by their British or American counterparts. However, they label this initial finding 'restrictive in its insightfulness' (p. 29) due to the fact that the quantitative results generated by larger corpora do not further an understanding of how or why hedges are used in face-to-face interaction.

Many corpus studies recommend that frequency analysis be complemented by a detailed consideration of the environment of key words through the use of concordances and collocational tools. For example, O'Keeffe and Adolphs (2008: 93) maintain that when there is a need to disambiguate form and function, corpus linguistics provides 'direct access to the source files and the exact location in the original conversations in which the items occurred.' For example, to add further insight into their raw frequency results, Farr and O'Keeffe (2002) explore the use of *would* as a hedging device in an Irish setting using two varietal sub-corpora from LCIE, a 55,000 word corpus of radio phone-in and a 52,000 word corpus of post-

observation teacher training interaction. Based on a qualitative examination of the hedges as they appear in context, in addition to confirming that Irish speakers soften face threatening acts such as disagreement or giving advice, they also found that very often speakers downtone when speaking about themselves, even where the propositional content is undisputed (e.g. *She would be eleven years dead now*). This led them to conclude that hedges have a broader pragmatic function in Irish English settings. They propose that in order to fully understand why speakers hedge it is necessary to consider the Irish socio-cultural context. They maintain that 'in Irish society, directness is very often avoided … "forwardness", which ranges from being direct to being self-promoting, is not valued' (Farr and O'Keeffe (2002: 42). Therefore, Irish speakers may feel added pressure to hedge in situations where British or American speakers may think it unnecessary. Farr and O'Keeffe's study demonstrates the merit of a two-pronged approach to the use of corpora in variational pragmatics, where intra-varietal, qualitative research involving smaller corpora is used to inform inter-varietal, quantitative corpus research.

Corpus-based variational pragmatic analysis can be further complemented by the demographic speaker information that accompanies conversations contained in many modern spoken corpora, thereby allowing both a micro- and macro-social interpretation of the corpus results. O'Keeffe and Adolphs (2008) analyse the form and function of response tokens across British and Irish English. To examine form, they analysed two one-million word corpus samples from CANCODE and LCIE. From these samples, they generated word and cluster lists and these were manually cross-checked with transcripts using concordancing. They demonstrate that, in terms of overall frequency, listener response tokens are far more frequent in British English than in Irish English. In order to compare the data functionally, they analysed two 20,000 word

subcorpora of casual conversation taken from LCIE and CANCODE. The demographic information provided by CANCODE and LCIE allowed them to closely match their data in terms of gender, age, social relationship, socio-economic class and genre of discourse. Accordingly, in both subcorpora the participants were female university students in shared accommodation, that were close friends and of similar age (around 20). By controlling for macro-social categories of gender, age and socio-economic class, O'Keeffe and Adolphs were able to make an accurate generalisation across two varieties of the same language. They again found that listener response tokens were more frequent among the British participants. While there was a degree of variation in terms of response token forms, their analysis revealed no pragmatic variation in how the response tokens functioned across the two subcorpora.

Previous variational research into hedging in family discourse has shown how it is more frequent in other discourse contexts. From an intra-varietal perspective, Farr *et al.* (2004) analysed the occurrence of hedging across various contexts such as family discourse, teacher training feedback, service encounters and female friends chatting in LCIE. They found that the lowest instance of hedging occurred in service encounters where 'there is an existing social schema for the interaction within exogenous roles' (p. 16-17), which simultaneously allows maximum transactional efficiency and minimum threat to face. The next least hedged context was the family where hedging was approximately 33% less frequent than in radio phone-in and 50% less frequent than in teacher training feedback. Farr *et al.*'s findings are consistent with a previous study by this researcher (Clancy 2005) where the occurrences of eight hedges prominent in Irish English were compared across two distinct context-types – family discourse and radio phone-in. It was found that hedges occur more than twice as frequently in radio phone-

in than in family discourse and this was attributed to the unique nature of family discourse. For example, some hedges, such as *kind of/sort of*, function to reduce the social distance between speakers and also to indicate the speaker's desire for a relaxed relationship with the addressee (Holmes 1993). These interpersonal aspects have to be worked at in contexts such as radio phone-in in order to create the pseudo-intimacy crucial to the success of the interaction (O'Keeffe 2006), however, this work is unnecessary in the family as the speakers perceive social distance as being negligible. Orpin (2005) maintains that corpus analysis allows the researcher to gain an insight into the semantic, connotative and prosodic meanings of a word and thereby enables the detailed construction of a word's 'semantic profile'. Similarly, the synergy of the variational pragmatic research agenda with a corpus linguistic methodology allows those working in variational pragmatics to construct a detailed 'pragmatic profile' of individual words, clusters or acts. This profile encompasses the social, cultural and discoursal information that influences a particular linguistic choice.

Using corpora for pragmatic research is, however, limiting on some levels. There are obviously some aspects of pragmatic analysis that are more suited to corpus analysis than others. For example, Jautz (2008: 147) observes that 'it is difficult, for instance, to investigate phenomena above the level of the word or phrase in corpora…Since corpora are not (yet) tagged for speech acts, it is not possible to search for all instances of gratitude in a speech act theoretical sense.' This particular aspect can, however, be overcome by the use of small corpora such as those in the case study presented here. Another issue particular to most spoken corpora is that transcripts are a written representation of a spoken text and are characterised by a tension between accuracy, readability and political issues of representation (see for example, Roberts

1997, Bird 2005). However, the continued development of multi-modal corpora will ensure that, in the future, the researcher will be able to align transcription with its audio-visual context. The comparability of corpora across language varieties can also be an issue because of the differing design criteria used in the construction of different corpora. There have been attempts within corpus linguistics to address this. For example, the International Corpus of English is designed for comparability across different varieties and the design framework for LCIE is based on the CANCODE matrix (see McCarthy 1998). Despite these issues, as the case study below will further illustrate, corpus linguistics affords the researcher access to (often large amounts) of naturally-occurring text, the ability to explore both form and function and the background information necessary to control for various macro- and micro-social factors, thereby providing variational pragmatics with a very compatible methodological tool.

1.2 The case study data

The two corpora represent spoken language collected in the home/family environment. Both families are from the Limerick City area in Ireland – one family is middle class from 'mainstream' Irish culture and one from the Irish Traveller Community. The Irish Traveller Community comprises a distinct cultural group that exists within Irish society. According to the Irish Central Statistics Office (2007), there are around 22,400 Travellers living in Ireland, accounting for just over 0.5% of the population. There are also Irish Travellers living in Northern Ireland and Britain (approximately 7,000) and the United States (approximately 10,000). Irish Travellers share a common ancestry and have shared fundamental cultural values that differ from those of the settled community and have a language of their own. Nomadism and the family are

core values of Traveller culture. Gmelch (1989) maintains that the Traveller family is the basic structural unit, as well as the primary unit of production and consumption. Maintaining family ties and ensuring contact with the extended family are fundamental to the Traveller way of life and this very often requires travel. As McDonagh (2000: 31) points out:

> It's important to remember that within Traveller society you have a mother, father and children, but that is not as important as the family group …You have one extended family and this is not seen in geographical terms. Settled people organise themselves within parishes and districts. Travellers organise within families.

Tovey and Share (2003: 472-473) claim that for Travellers, nomadism is a more significant marker of ethnicity than language and it has emerged as 'their most important distinctive attribute.' There is, however, increasing evidence that Travellers now view their language, *Cant*, as a central symbol of their ethnicity [3]. For example, one of the predecessors of the Irish Traveller Movement was called *Minceir Misli* (*Us Travellers*), and more recently, Travellers have chosen the word *pavee* to refer to themselves.

As Table 2 shows, the middle class family corpus (SettCorp) consists of one hour of audio recordings. The total number of members of the family is six (two parents and four siblings), however, not all recordings feature six speakers. SettCorp, in addition to conversations featuring both parents and children, also contains instances of conversation that feature the siblings in interaction with one another in the absence of the parents. The Traveller family corpus (TravCorp) is composed of forty-five minutes of audio recordings. The total number of members

of the Traveller family is eight (two parents and six children). All conversations feature at least one parent in conversation with his/her children.

Table 2: Description of the two datasets

	SettCorp	TravCorp
Length of recording	60 minutes	45 minutes
Number of speakers	6	8
Number of words	12531	3172

In relation to the participants, detailed demographic information was collected for both families. The gender profile for both families is the same with equal numbers of male and female participants. In both corpora, the recordings were restricted to the immediate family and the home environment – in the case of the settled family, a house, and for the Traveller family, a mobile home. All names have been anonymised and pseudonyms given, and any references that could identify the exact location of the recordings have been removed. Apart from these changes, that data remains uncensored. Therefore, the data collected is naturally occurring, spontaneous, casual conversation. In relation to the extracts featured in this paper, all transcriptions are contextualised and the role of the speaker in the family, for example, father or daughter, provided at the beginning of each speaker turn. The extracts are marked SC for those taken from SettCorp

and TC for those taken from TravCorp. Where there are extracts that feature two or more daughters or sons, these are labelled <Daughter 1>, <Daughter 2> etc. Other information relevant to the extracts such as speaker age is given where necessary.

The focus of the findings from the case study is the variational distribution of the occurrence of hedges in both SettCorp and TravCorp. According to Brown and Levinson (1987: 145 [original emphasis]) a hedge:

> …is a particle, word, or phrase that modifies the degree of membership of a predicate or noun phrase in a set; it says of that membership that it is *partial*, or true only in certain respects, or that it is *more* true and complete than perhaps might be expected.

For example, hedges such as *I think* (labelled a *quality hedge* by Brown and Levinson 1987: 164) allow the speaker to avoid full responsibility for the truth of his/her utterance, distancing both her/himself and the hearer from the act, thereby satisfying or redressing the hearer's negative face. Therefore, hedges downtone the illocutionary force of an utterance allowing the speaker to weaken his/her commitment to its propositional content. Hedges have a lesser role to play in positive politeness: linguistic actions aimed at building on indices of solidarity such as in-group membership, modifying extremes on the value scale such as *beautiful* or *revolting*. Therefore, Brown and Levinson (1987: 116-117) claim that in the utterance *It's really beautiful, in a way*, the hedge *in a way* allows the speaker to avoid the precise communication of his/her attitude, 'leaving it up to the addressee to figure out how to interpret it'. They maintain that by using one of these hedges, the speaker calls upon the hearer to use the common knowledge between them to interpret speaker attitude thereby appealing to the hearer's positive face.

1.3 Findings

Schneider and Barron's (2008) research agenda for variational pragmatics specifies an initial concentration on the impact of macro-social factors on pragmatic variation. To date, the majority of corpus linguistic research into variational pragmatics has focused on regional variation across national varieties of a language, such as between American and British English (see for example, Tottie 1991, McCarthy 2002). This case study aims to contribute to the pragmatic characterisation of Irish English. Furthermore, as outlined in Section 1.2, the corpora in this case study were collected in such a way as to ensure that region and gender are comparable. This serves to demonstrate the impact of age and/or social class and/or ethnicity on pragmatic variation − in particular, in relation to the occurrence of hedges. In order to find items with the *potential* to function as hedges in family discourse, a two-pronged approach was taken. Firstly, frequency lists were generated in both SettCorp and TravCorp based on the ten most frequent single-word and two-word hedges in LCIE (identified by Farr *et al*., 2004), the results of which are presented in Table 3:

Table 3: Items with the *potential* to function as hedges in SettCorp and TravCorp

Item	Frequency SettCorp	Frequency TravCorp	Total
like	83	9	**92**
just	39	2	**41**
you know	24	4	**28**

I think	18	0	**18**
actually	16	0	**16**
really	16	0	**16**
a bit	10	2	**12**
probably	9	0	**9**
kind of/sort of	1	0	**1**
I suppose	1	0	**1**
Total	**222**	**17**	**239**

Table 3 demonstrates that in SettCorp, items with the potential to hedge appear to occur far more frequently in this corpus than in TravCorp. However, because of the disparity in size between the two corpora, it is necessary to normalise the figures. When normalised, items with the potential to hedge occur with a frequency of 177 instances per 10,000 words in SettCorp and 53 instances per 10,000 words in TravCorp. In order to perform a functional analysis, concordances were used to exclude all non-hedging instances of the top five markers listed in Table 3. Therefore, Table 4 illustrates the actual number of instances of hedging for the markers *like*, *I think*, *just*, *you know* and *actually*:

Table 4: Frequency of occurrence of five hedges across SettCorp and TravCorp

Hedge	SettCorp	TravCorp
like	19	0
I think	16	0

just	14	1
you know	13	2
actually	11	0
Total	**73**	**3**

Table 4 clearly shows that these five hedges have a far higher frequency in SettCorp than in TravCorp. Again, due to disparity in corpus size, it is necessary to normalise these figures. When normalised, these hedges occur at a rate of 58 times per 10,000 in SettCorp and 9.5 times per 10,000 in TravCorp. This suggests that the five top hedges in Irish English are six times more frequent in the everyday speech of the settled family than the Traveller family. In order to construct a 'pragmatic profile' of these hedges, they are compared individually by frequency in Figure 1:

Figure 1: Frequency of items functioning as hedges in SettCorp and TravCorp (normalised per 10,000 words)
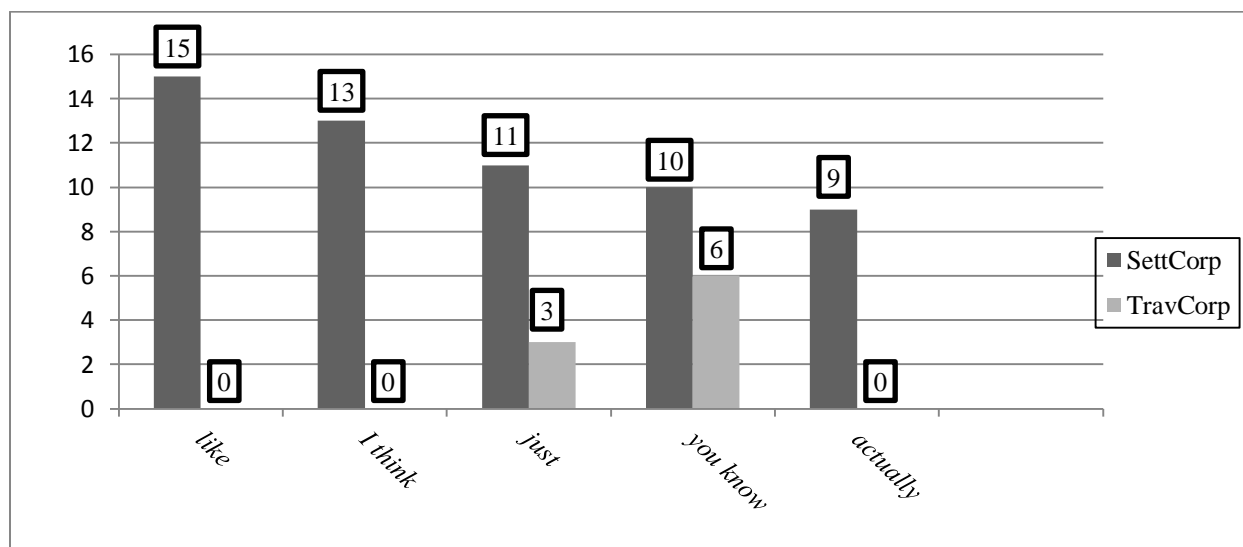

17

Table 4 and Figure 1 demonstrate that hedging appears to be relatively rarely used in the Traveller Community in comparison to the settled community. The two most common hedges in SettCorp, *like* and *I think*, do not feature in TravCorp, despite the comparability of the corpora in terms of region, gender and age. Similarly, *actually* does not occur in TravCorp, however, this marker has a frequency of almost nine occurrences per 10,000 words in SettCorp. The marker *just* is more than three times more frequent in SettCorp than in TravCorp. The only marker with a comparable frequency is *you know*, six and ten occurrences per 10,000 words in TravCorp and SettCorp respectively. In order to account for the pattern of variation in the occurrences of these hedges in the two families, it is necessary to discuss the differing influences of macro-social factors such as ethnicity and social class on the two corpora.

1.4 Discussion

The Traveller Community exhibits some of the characteristics of East Asian collectivist cultures such as the primacy of the family unit and also the hierarchies that exist within it. Gormally (2005: 79) attests to the importance of position in Traveller families noting that 'all children who mentioned family members were able to account for their position within the family – for one child that meant knowing that he was the second youngest of a family of twenty-three.' An examination of collectivist discourse styles led Scollon and Scollon (1995: 131) to maintain that 'individual members of a culture are not seen as independently acting individuals, but rather they are seen as acting within hierarchies of kinship and other relationships.' Extract (1 TC) features members of the Traveller family at the breakfast table. The father is attempting to get the youngest member of the family to finish his breakfast, all the other children have finished:

(1 TC)

\<Father\>	C'mon eat the breakfast **baby**.
\<Son 1\>	How you doin' Johnny?
\<Son 2\>	They're daddy's shoes.
\<Father\>	They're daddy's shoes are they **son**?

The father's use of child-specific kin titles such as *baby* and *son* (in bold) could be interpreted as downplaying the value of autonomy evident in a full first name, the emphasis is instead on belonging and interdependence (cf. Blum-Kulka, 1997). This may provide evidence of the close social networks that exist within the Irish Traveller Community. These kinship networks are based around family and extended family and this unit also provides Travellers with both their social and work groupings. Therefore, their primary relationships in the family are the same as their secondary relationships in the workplace, a trait common in other marginalised communities in the English-speaking world (see Youmans 2001). In this kinship culture, the importance of the family unit, and one's position in it, is reinforced by the use of these kin titles. According to Markkanen and Schröder (1997: 8), 'the surer a speaker feels about his or her position vis-à-vis the interlocutor, the less need there is for hedging for the purposes of self-protection.' Therefore, arguably, the assuredness of their position in the family reduces the need for Traveller family members to use hedges.

On the other hand, the individualistic nature of settled culture involves a recognition of social autonomy and independence. In contrast to the Traveller Community, members of the settled community move outside the family unit in order to enter the workplace and establish extended social networks. The settled parents are aware that the children will move from the family into the educational or work sphere and by hedging in their speech, are equipping the children with the necessary tools to do so. In addition to this, because the family are using hedges in other 'external' speech situations, it is manifest in their talk when they return to the family unit. In extract (2 SC), the siblings, in the absence of the parents, are gossiping about the physical appearance of a student enrolled on the same university course as Daughter 1, a subject that is considered a sensitive one in many cultures:

(2 SC)

<Daughter 1>   He wasn't outside today. He'd actually give it to you.

<Son 1>   Fat boy.

<Daughter 1>   He's fierce healthy now I'm not jokin' you [4].

<Son 1>   He's fierce fat too.

<Daughter 1>   He's not actually that heavy.

<Son 2>   Are you callin' people fat?

<Daughter 1>   **I think** he was though the year before that I do.

```
                                              [
<Daughter 2>                    Conor in relation to you every one is fat.


<Daughter 1>    No but **I think** he was heavy before.

                                              [
<Son 2>                         God you're awful mean you skin head knacker[5].


<Son 1>         I am getting fat though.


<Daughter 1>    **I think** he was heavier before.


<Son 2>         I heard Jennifer was sayin' that and all here look he's getting fat.

                                      [
<Daughter 2>                          He's getting fat.


<Son 1>         I'm puttin' on weight.
```

In his first utterance, Son 1 asserts that this student is a *fat boy* and Daughter 1 appears to contradict him using a series of hedged lexical reformulations centred around the word *fat*, for example, *he's fierce healthy* or *he was heavier before*. In addition, as the conversation progresses, Daughter 1 appears to realise that her opinion is different to Son 1 and she reformulates her position using *I think* (in bold) on three occasions and this functions to soften her disagreeing acts thus protecting her face and how she is perceived within the sibling group. Murphy (2010) analysed the casual conversation of females in their twenties and discovered a high level of hedging in order to cover uncertainty or downtone assertiveness in the event that the speaker is wrong or his/her opinion differs from that of the other interlocutors in the group. In addition, in

this case, it may be that the siblings, especially the female ones, have acquired the cultural knowledge that hedging is necessary around sensitive issues and that 'weight' is one of these.

The pattern of pragmatic variation presented here for TravCorp and SettCorp could also be attributable to a macro-social factor strongly linked to socio-economic status, that of educational achievement. The demographic information for both families reveals a noticeable difference in educational attainment. In SettCorp, three of the children are students at third level (two of whom are postgraduate students) and one is a student at second level. Although there are no educational qualifications recorded for the children in TravCorp, in the Traveller Community as a whole, two-thirds of all school leavers are educated to, at most, primary level (Irish Central Statistics Office 2004). In addition to this, a study into the educational background of Travellers in Galway, a city in the west of Ireland, revealed that in contrast with a rate of 26% in the settled population, no Traveller had a third level degree (Irwin, 2006). Markkanen and Schröder (1997: 9), through an analysis of hedging in academic writing, claim that hedges acquire their meaning 'through a process of author-reader interaction, on the basis of the text and the communicative situation.' This interaction is somewhat controlled by *culture*, 'since people who belong to a particular language community normally shared socially determined aesthetic ideals through their shared educational background' (*ibid.*). Brown and Levinson (1987: 250) argue that English-speaking academic speech communities constitute 'negative politeness cultures', which, according to Holmes (1984: 348) demonstrate a 'fascination for devices which attenuate negatively affective speech acts.' In transcending the academic sphere, Youmans (2001) notes a preference for negative politeness in the everyday life of educated Anglos, middle-class speakers representative of mainstream US society. Therefore, in contrast to the Traveller Community,

markers of negative politeness like hedges have been acquired by the members of the settled family in the educational sphere, and their usage has then been invoked in the family setting. O'Sullivan (2004), in a study of the accommodative phenomena of teenage Traveller girls in secondary education, illustrates that they use *like* as a hedge 2.5 times more frequently in interview settings than in informal conversation. As the interviews were performed by a member of the settled community, she claims that this indicates a desire on the part of Traveller girls to conform to the speech norms of their settled peers and, in doing so, gain social acceptance. In informal conversation with each other and with no member of the settled community present, the girls use less hedging.

Huspek (1989), seeking to account for instances of linguistic variability and power, analyses occurrences of *you know*/*I think* in American industrial workers' speech, a group he delineates as socially disadvantaged due to their occupation and educational qualifications. He notes that among the workers, the ratio of occurrence of *you know* to *I think* is 8:1. He also observes 'strong sentiments against the use of the latter sequence [*I think*]' (p. 670). Accordingly, all instances of *I think* (except one) are used in conjunction with modals and the markers *you know* and *I don't know* by the workers. This, he claims, allows the workers to express individual opinions while showing consideration for the group, necessary because in the workplace it is the group that wields the most power in the form of union activity etc. Therefore, in order to oppose the dominant educational ideology, the Travellers as a group may have created what Huspek terms 'verbal resistance stratagems' (p. 681) that operate to challenge disadvantage. If hedges are acquired in the educational sphere, it could be claimed that the Traveller Community have rejected these hedges due to a past Traveller education policy that viewed education as a matter

of settlement, a way of taking the Traveller out of the child (Pavee Point Travellers Centre 2009). It is also possible that the Traveller Community's non-use of hedges could have repercussions for them when they enter mainstream culture. Youmans (2001) equates the Anglo use of *I think* with the language required for success in wider society. She contends that the corresponding refusal by Chicano speakers to match the dominant class' language norms perpetuates their position as a non-powerful, disadvantaged group in American society. Akin to the Chicanos, the Traveller Community's 'failure' to employ linguistic forms and functions such as the use of *I think* for hedging purposes, although strengthening the bonds within their own cultural grouping, may have a direct influence on their continuing marginalisation in modern-day Ireland.

The macro-socio factors of age, ethnicity and socio-economic status can be seen to account for both the presence and absence of hedges within the two corpora. As has been discussed, the Traveller Community place family at the centre of their society and maintaining family ties and ensuring contact with the extended family are fundamental to the Traveller way of life. Combining the strength of family ties to the fact that when Travellers 'travel', they do so in order to ensure sustained contact with the extended family, leads to a system of social networks unique to this culture in Irish society. On the other hand, the Irish middle class, although bereft of 'ethnic' status, could be said to be distinct from other socio-economic groupings in Irish society due to, for example, high levels of educational achievement and high social mobility. It is these distinguishing characteristics that account for the fact that the hedges *like*, *I think*, *just*, *you know* and *actually* are far more frequent in SettCorp than in TravCorp. Indeed, it could be said that these hedges represent those that are critical to politeness in 'mainstream' Irish culture. They are the absolute minimum needed for polite interaction among
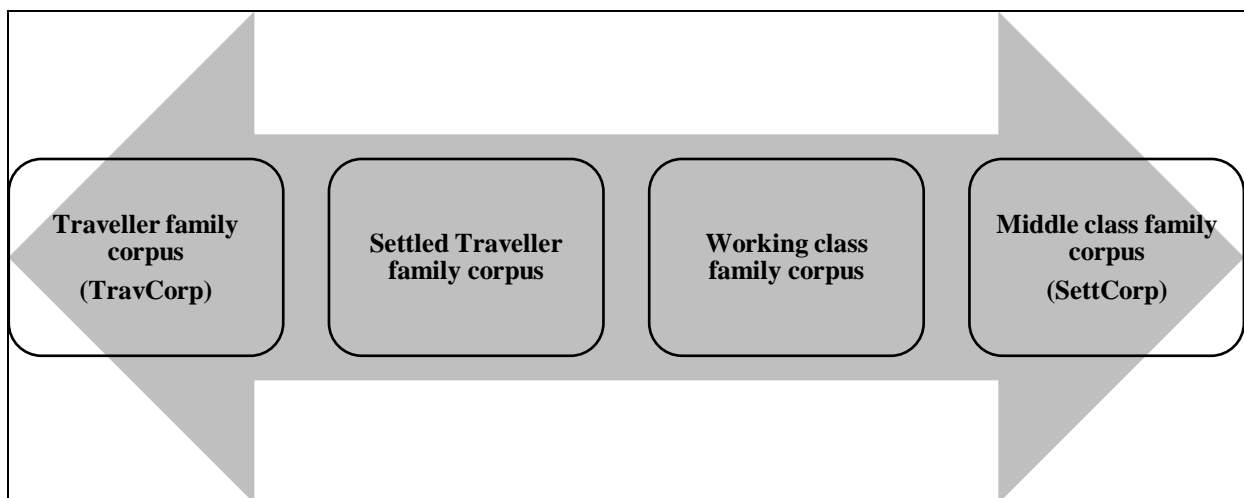
participants in Irish society and ensure a smooth transition from the family community of practice to the wider social world. Correspondingly, they are in a sense 'redundant' in the Traveller Community given that they rarely move into the realm of mainstream society.

1.5 Conclusion

The case study presented here does not make any claims that the two families featured are representative of their respective communities. This, coupled with the assertion that differences in the two families' pragmatic systems are due to macro-social variables such as ethnicity and socio-economic status, points towards the primary avenue for extension of the study. In order to bridge the social and ethnic divide between TravCorp and SettCorp, a logical first step would be to build corpora that would connect them, as illustrated in Figure 2:

Figure 2: Bridging the social and ethnic divide between TravCorp and SettCorp[1]

Therefore, as Figure 2 demonstrates, it is proposed that future research construct both a settled Traveller family and working class family corpus that are broadly comparable to TravCorp and SettCorp. This could potentially allow for the consolidation of the findings from the present study.

Corpus linguistics and variational pragmatics have been successfully blended in a number of studies: corpora and the corpus linguistic tools that allow researchers to mine them provide an empirical bent for variational pragmatic research. The variational pragmatic research agenda facilitates the use of corpora in the study of pragmatic variation between different varieties of a language and between different groups of speakers. For example, the case study presented demonstrates how applying corpus tools to specific, situated speech contexts – in this case family discourse – can result in an intra-varietal appraisal of pragmatic norms between different two cultures. It may well be that the study of the pragmatic practices of two Irish families with different social and ethnic backgrounds could contribute in some way towards understanding any linguistic misconceptions that may be held either by settled people about Travellers or vice versa. The paper also highlights the importance of small corpora in variational pragmatic research. Small corpora, similar to those presented in the case study, are relatively easily assembled and analysed which results in 'current' linguistic knowledge. Small, register-specific corpora also afford the opportunity to examine nuances in pragmatic use rather than simply seeking to formulate generalisations. While there are undoubtedly some difficulties to be overcome when using corpora to study variational pragmatics, the benefits far outweigh any misgivings. Schneider and Barron (2008) maintain that the fields of dialectology and pragmatics are akin to

fiancées that should get married quickly (they also specify that they should have many healthy children) – this paper suggests that corpus linguistics should conduct the marriage ceremony.

Notes

---

[1] The Limerick Corpus of Irish English (LCIE) is a one million word corpus of naturally occurring spoken Irish English (for more details see Farr *et al.* 2004). Both frequency lists are unlemmatised. The spoken BNC frequency list is taken from Leech *et al.*, 2001.

[2] Both frequency lists are unlemmatised. The spoken BNC frequency list is taken from Leech *et al.*, 2001.

Table 2: Description of the two datasets

	SettCorp	**TravCorp**
Length of recording	60 minutes	45 minutes
Number of speakers	6	8
Number of words	12531	3172

Table 3: Items with the *potential* to function as hedges in SettCorp and TravCorp

Item	Frequency SettCorp	Frequency TravCorp	Total
like	83	9	**92**
just	39	2	**41**
you know	24	4	**28**
I think	18	0	**18**
actually	16	0	**16**
really	16	0	**16**
a bit	10	2	**12**
probably	9	0	**9**
kind of/sort of	1	0	**1**
I suppose	1	0	**1**
Total	**222**	**17**	**239**

Table 4: Frequency of occurrence of five hedges across SettCorp and TravCorp

Hedge	SettCorp	TravCorp
like	19	0
I think	16	0
just	14	1
you know	13	2
actually	11	0
Total	**73**	**3**

Figure 1: Frequency of items functioning as hedges in SettCorp and TravCorp (normalised per 10,000 words)
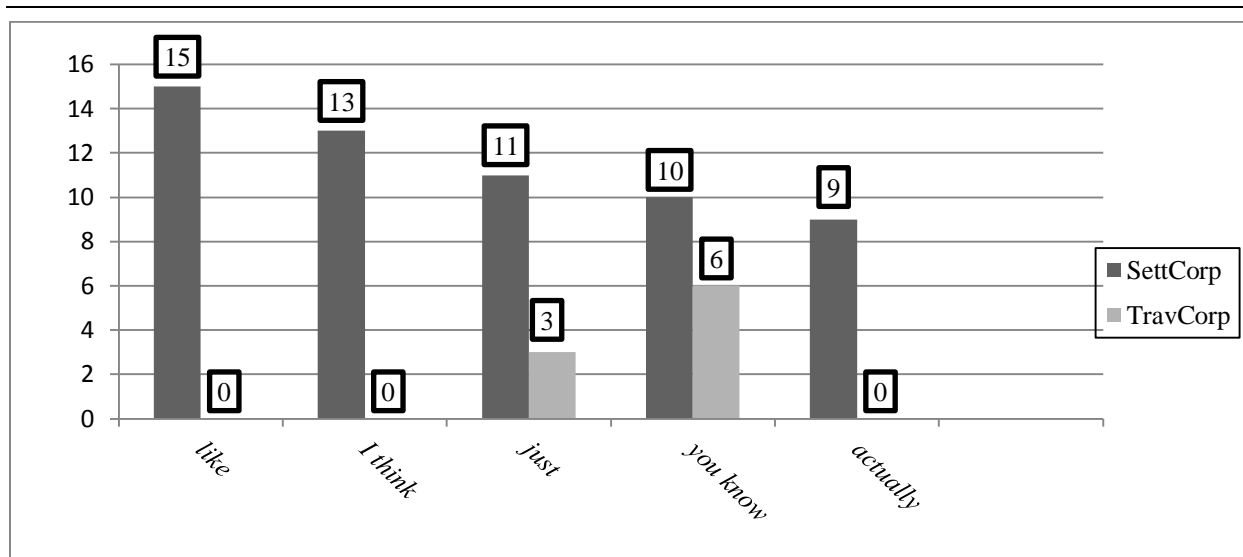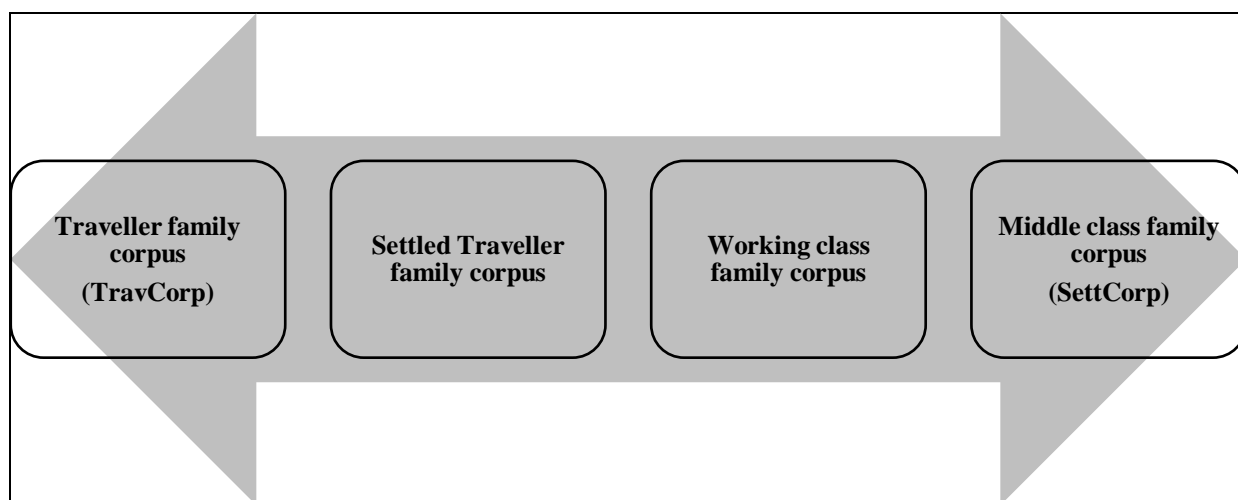
Figure 2: Bridging the social and ethnic divide between TravCorp and SettCorp[2]

[3] The academic name for the language spoken by Travellers is *Shelta*, but Travellers themselves refer to it as *Gammon* or *Cant*.

[4] *Fierce* is often used in Irish English to mean *very*.

[5] *Knacker* (/nækər/) is a derogatory term in Irish English typically used to describe people from low-income, working class backgrounds who engage in anti-social behaviour. Interestingly, the word originated as a derogatory reference to members of the Travelling Community.

References

Atkins, S., J. Clear and Ostler, N. 1992. "Corpus design criteria." *Literary and Linguistic Computing*, 7(1), 1-16.

Barron, A. and Schneider, K. 2009. "Variational pragmatics: Studying the impact of social factors on language use in interaction." *Intercultural Pragmatics*, 6(4), 425-442.

Biber, D., 1993. "Representativeness in corpus design." *Literary and Linguistic Computing*, 8(4), 243-257.

Bird, C., 2005. "How I stopped dreading and learned to love transcription." *Qualitative Inquiry*, 11(2), 226-248.

Blum-Kulka, S. 1997. *Dinner Talk: Cultural Patterns of Sociability and Socialisation in Family Discourse*, New Jersey: Lawrence Erlbaum.

Brown, P. and Levinson, S. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

Central Statistics Office, Ireland 2004. *Census 2002: Volume 8 – Irish Traveller Community*. Dublin: Stationery Office.

Central Statistics Office, Ireland. 2007. *Census 2006: Volume 5 – Ethnic or Cultural Background (Including Irish Travelling Community)*. Dublin: Stationery Office.

Clancy, B. 2005. "*You're fat. You'll eat them all*: Politeness strategies in family discourse". In: A. Barron & K. Schneider (Eds.), *The Pragmatics of Irish English*. Berlin: Mouton de Gruyter, 177-199.

Clear, J. 1992. "Corpus sampling". In: G. Leitner (Ed.), *New Directions in English Language Corpus Methodology*. Berlin: Mouton de Gruyter, 21-31.

Crowdy, S. 1993. "Spoken corpus design." *Literary and Linguistic Computing*, 8(4), 259-265.

Farr, F. and O'Keeffe A. 2002. "*Would* as a hedging device in an Irish context: An intra-varietal comparison of institutionalised spoken interaction". In: R. Reppen, S. Fitzmaurice & D. Biber (Eds.), *Using Corpora to Explore Linguistic Variation*. Amsterdam: John Benjamins, 25-48.

Farr, F., Murphy, B. and O'Keeffe, A. 2004. "The Limerick Corpus of Irish English: Design, description and application." *Teanga*, 21, 5-30.

Gmelch, S. 1989. "From poverty subculture to political lobby: The Travellers' rights movement in Ireland". In: C. Curtin and T. Wilson (Eds.), *Ireland from Below: Social Change and Local Communities*. Galway: Galway University Press, 301-320.

Gormally, E. 2005. *Whiddin to the Gauras: Talking to our own*. Dublin: Veritas.

Holmes, J. 1984. "Modifying illocutionary force." *Journal of Pragmatics*, 8, 345-365.

Holmes, J. 1993. "New Zealand women are good to talk to: An analysis of politeness strategies in interaction." *Journal of Pragmatics*, 20, 91-116.

Hunston, S. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Huspek, M. 1989. "Linguistic variability and power: An analysis of *you know/I think* variation in working class speech." *Journal of Pragmatics*, 13, 661-683.

Irwin, A. 2006. *Stall Anoishe! Minceirs Whidin: Analysing the Reality for Travellers in Galway City*. Galway Traveller Movement.

Jautz, S. 2008. "Gratitude in British and New Zealand radio programmes: Nothing but gushing?" In: K. Schneider & A. Barron (Eds.), *Variational Pragmatics: A Focus on Regional Varieties in Pluricentric Languages*. Amsterdam: John Benjamins, 141-178.

Leech, G., Rayon, P. and Wilson, A. 2001. Companion website for: *Word Frequencies in Written and Spoken English: based on the British National Corpus*. available on-line at http://ucrel.lancs.ac.uk/bncfreq/ [accessed 09-12-2009].

Markkanen, R. and Schröder, H. 1997. "Hedging: A challenge for pragmatics and discourse analysis". In: R. Markkanen & H. Schröder (Eds.), *Hedging and Discourse: Approaches to the Analysis of a Pragmatic Phenomenon in Academic Texts*. Berlin: Mouton de Gruyter, 3-18.

McCarthy, M. 1998. *Spoken Language and Applied Linguistics*. Cambridge: Cambridge University Press.

McCarthy, M. 2002. "Good listenership made plain: British and American non-minimal response tokens in everyday communication". In: R. Reppen, S. Fitzmaurice & D. Biber (Eds.), *Using Corpora to Explore Linguistic Variation*. Amsterdam: John Benjamins, 49-72.

McDonagh, M. 2000. "Ethnicity and culture". In: F. Murphy & C. McDonagh (Eds.), *Travellers: Citizens of Ireland*. Dublin: Parish of the Travelling People, 26-31.

McEnery, T., Xiao, R. and Tono, Y. 2006. *Corpus-Based Language Studies: An Advanced Resource Book*. London: Routledge.

Murphy, B. 2010. *Corpus and Sociolinguistics: Investigating Age and Gender in Female Talk*. Amsterdam: John Benjamins.

O'Keeffe, A. 2006. *Investigating Media Discourse*. London: Routledge.

O'Keeffe, A. and Adolphs, S. 2008. "Response tokens in British and Irish discourse: Corpus, context and variational pragmatics". In: K. Schneider & A. Barron (Eds.), *Variational Pragmatics: A Focus on Regional Varieties in Pluricentric Languages*. Amsterdam: John Benjamins, 69-98.

Orpin, D. 2005. "Corpus linguistics and critical discourse analysis: Examining the ideology of sleaze." *International Journal of Corpus Linguistics*, 10(1), 37-61.

O'Sullivan, J. 2004. *Accommodative phenomena in the speech of Traveller girls: A corpus-based analysis of the convergences to the speech patterns of the settled community*. Unpublished MA thesis, University of Limerick.

Pavee Point Travellers' Centre, Ireland, 2009. *Education and Cultural Rights Fact Sheet*, available on-line at http://www.paveepoint.ie/pdf/FactSheet-education.pdf [accessed 02-09-2009].

Plevoets, K., Speelman, D. and Geeraerts, D. 2008. "The distribution of T/V pronouns in Netherlandic and Belgian Dutch". In: K. Schneider & A. Barron (Eds.), *Variational Pragmatics: A Focus on Regional Varieties in Pluricentric Languages*. Amsterdam: John Benjamins, 181-210.

Roberts, C. 1997. "Transcribing talk: Issues of representation." *TESOL Quarterly*, 31(1), 167-172.

Schneider, K., and Barron, A. 2005. "Variational pragmatics: Contours of a new discipline". Paper presented at the *9th International Pragmatics Conference*, Riva del Garda, July 10-15.

Schneider, K., and Barron, A. 2008. "Where pragmatics and dialectology meet: Introducing variational pragmatics". In: K. Schneider & A. Barron (Eds.), *Variational Pragmatics: A Focus on Regional Varieties in Pluricentric Languages*. Amsterdam: John Benjamins, 1-32.

Scollon, R. and Scollon, S. 1995. *Intercultural Communication: A Discourse Approach*. Oxford: Blackwell.

Tognini-Bonelli, E. 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamins.

Tottie, G. 1991. "Conversational style in British and American English: The case of backchannels". In: K. Aijmer & B. Altenberg (Eds.), *English Corpus Linguistics*. London: Longman, 254-271.

Tovey, H. and Share, P. 2003. *A Sociology of Ireland*. Dublin: Gill and Macmillan.

Youmans, M. 2001. "Cross-cultural differences in polite epistemic modal use in American English." *Journal of Multilingual and Multicultural Development*, 22(1), 57-73.

Wolfram, W. and Schilling-Estes, N. 2006. *American English: Dialect and Variation*. London: Blackwell.

Author's Address

Brian Clancy

Mary Immaculate College

University of Limerick

South Circular Road

Ireland

Brian.Clancy@mic.ul.ie