WILLEM HOLLMANN AND ANNA SIEWIERSKA

Corpora and (the need for) other methods in a study of Lancashire dialect

Abstract: This paper is based on a nascent project on Lancashire dialect grammar, which aims to describe the relevant features of this dialect and to engage with related theoretical and methodological debates. We show how corpora allow one to arrive at more precise descriptions of the data than was previously possible. But we also draw attention to the need for other methods, in particular modern elicitation tasks and attitude questionnaires developed in perceptual dialectology. Combining these methods promises to provide more insight into both more general theoretical issues and the exact nature of the object of study, namely Lancashire dialect.

1. Introduction

This paper describes a nascent project on the grammar of Lancashire dialect. In line with the theme of this issue we focus on methodological issues, in particular matters concerning the use of corpora, both advantages and challenges. Section 2 presents some background information about the aims of the project, situating it within the recent surge in interest in dialect grammar. Section 3 offers an illustration of the insights that may be obtained from using corpora as well as the limitations of corpus studies, with reference to three ditransitive constructions. Section 4 shifts the stage to two other variables we have started to analyse, namely agreement in past tense BE and the presence, reduction or omission of the definite article. We will argue for the use of an additional data collection method, viz. modern elicitation tasks (cf. e.g. Cornips and Jongenburger 2001, Cornips and Poletto 2005). Section 5 wraps up the discussion with an outlook and some suggestions for possible directions of future research such as perceptual dialectology (cf. e.g. Preston 1989, 1999). In view of the problems associated with restricting oneself to one methodology only, we conclude that a multi-method approach is the best way forward.

2. Background and aims of the project

The study of dialect grammar lies at the intersection of two areas of linguistics: dialectology and theoretical linguistics. For partly historical reasons researchers have

traditionally avoided this intersection. Gerritsen (1991, 9) reports that only about 5 per cent of the maps in dialect atlases are devoted to syntax. However, recent developments such as the very large-scale project on the Syntactic Atlas of Dutch Dialects[1] and a number of publications in this area, such as the volume edited by Kortmann et al. (2005), show that interest is now on the increase.

The reasons behind the traditional neglect of dialect grammar are several. First, syntactic areas tend to be much larger (i.e. geographically speaking) than phonological ones, making them much harder to identify (see e.g. Bucheli and Glaser 2002:46). Second, the Generative paradigm, which has dominated syntactic research in the last four decades, has traditionally not been concerned with variation, including regional variation. And while the importance of syntactic variation has been recognised within the rival functional-cognitive paradigm, the existence of such variation has not been incorporated explicitly and/or adequately into the actual models of grammar developed, with the partial exception of Halliday's Systemic Functional Grammar (cf. Butler 2003). A fourth factor underlying the lack of scholarly attention given to dialect grammar relates to the availability of the data. Needless to say, there are many more grammatical patterns than there are phonemes. Accordingly, research into syntactic variation requires vast amounts of transcribed data, incomparably more than research into phonetic or phonological variation.

Several recent developments have counterbalanced the above and paved the way for research on the grammar of dialects. Advances in corpus linguistics have made it possible to access sufficiently large amounts of data to enable the study of grammatical phenomena. Both dialectologists and theoretical linguists have responded to this, as evidenced by e.g. Tagliamonte and Lawrence (2000) and the papers in Kortmann et al. (2005). Considerably improved syntactic elicitation techniques, most notably those introduced within the context of the previously mentioned SAND project, have minimised well-known pitfalls, e.g. the Observer's Paradox and adjustment to the standard (Labov 1972), biases caused by pragmatic infelicities (Greenbaum 1973) and repetition effects (Bock 1986). The SAND researchers also draw on advanced research in the interpretation of the results from elicitation tasks, e.g. methods developed by Bard et al (1996), Schütze (1996) and Cowart (1997). Taking their cue from research in related disciplines such as psychology, these authors all advocate far more sophisticated methods of interpreting grammaticality judgements than had been the norm in linguistics, especially in Generative Grammar.

The Lancashire dialect grammar project described in this paper fits in with the larger trend towards exploring the interplay between linguistic theory and regional variation. Methodologically, it sets out to exploit the advances in corpus linguistics, whilst also recognising the value of modern elicitation techniques. In addition, we identify a third important recent research trend, namely perceptual dialectology. This strand, which is primarily associated with Dennis Preston (cf. 1989, 1999) starts

---

[1] Also known as SAND, see e.g. http://www.meertens.knaw.nl/ projecten/sand/sandeng.html (12 October 2005).

from the observation that dialects are not purely linguistic entities, but are to a large extent social constructs, and should be studied as such. The methodological implications of this for our project will briefly be addressed below.

The motivation for studying the Lancashire dialect is threefold. First, whereas relatively much work has been done on the grammar of regional varieties such as those of the South West (e.g. Ihalainen 1991, Wagner 2005), Tyneside (e.g. Beal and Corrigan 2002, Beal 2004) and Yorkshire (e.g. Tagliamonte 1998, Tagliamonte and Lawrence 2000), Lancashire is decidedly understudied. (Shorrocks' (1999) grammar of the Bolton area variety comes close, but after the Local Government Act, implemented in 1974, Bolton is no longer part of Lancashire.) Second, informal impressions suggest that whatever this lack of scholarly attention has been caused by, it is not an absence of interesting patterns. Third, given our institutional affiliation we have easy access to relevant data.

With respect to our aims, unsurprisingly we seek to redress the neglect of Lancashire grammar to some extent. But our project is far from descriptive only: in analysing interesting variables we hope to engage with more general debates of a theoretical nature. Our approach may be termed "organic" in that the theoretical debates we engage in by and large depend on the Lancashire features that we have set out to analyse. Thus our considerations of ditransitives in Section 3 bear on the interface between semantics and syntax, while the preliminary analyses of agreement in past tense BE and of definite article reduction in Section 4 have potential implications for our understanding of the notion of 'sociolinguistic salience' (cf. e.g. Kerswill and Williams 2002). The discussion throughout is framed so as to make it clear where corpora are helpful, and where there is a need for supplementary methods.

3. The value of corpora: ditransitive constructions

British English exhibits a particularly rich array of encoding possibilities of the theme and recipient in ditransitive clauses obtaining essentially, though perhaps not exclusively, when both the theme and recipient are pronominal. In addition to the typical prepositional construction as in (1a), there are two double object constructions, the first with the recipient preceding the theme, as in (1b), and the second with the theme preceding the recipient, as in (1c).

(1)   a. *John gave it to me.*
        b. *John gave me it.*
        c. *John gave it me.*

As documented in Siewierska and Hollmann (2005), there is considerable confusion in the literature as to the presence and distribution of these three constructions in British dialects. Moreover, virtually nothing is known about the nature of the ditransitive verbs which may occur in each of the three constructions and the nature of the person pronouns which they permit as themes and recipients.

To shed some light on the issue we investigated the properties of the three ditransitive constructions in four corpora of Lancashire dialect, namely the Lancashire part of the British National Corpus,[2] the Lancashire texts of the Survey of English Dialects Incidental Recordings Corpus,[3] the Lancashire section of the Freiburg English Dialect Corpus[4] and the Lancashire part of the Helsinki Corpus of British English Dialects.[5] These corpora, numbering in all about 500,000 words, yielded 449 ditransitive clauses, of which only 77 (17%) featured both pronominal recipients and themes. This figure dropped to 43 (10%) when nonspecific pronominal recipients or themes such as *someone*, *one*, *any*, *some*, *something* were disregarded. Nonetheless, even among this small number of relevant ditransitive clauses with themes and recipients rendered by personal pronouns there were instances of all three ditransitive constructions, 20 prepositional ones and 23 double object ones. Importantly, there were more instances of the double object construction with theme preceding the recipient (15, i.e. 65%) than with the recipient preceding the theme (8, i.e. 35%). Moreover, the theme-before-recipient order in the double object pattern occurred not only with the verb *give* but also with the verbs *send*, *offer* and *show*.

The above findings seem to be corroborated by the data stemming from yet another corpus, which we are compiling from transcripts based on recordings from the North West Sound Archive.[6] It contains almost 30 hours of transcribed speech. The results obtained from the NWSA are preliminary, however, since the transcripts must be subjected to a thorough revision, a process which has only just begun.

The small investigation outlined above provided the first piece of evidence that there is at least one British dialect in which the placement of the theme before the recipient may well be the preferred pattern in the double object construction with pronominal themes and recipients. It also revealed that this pattern is not confined to the verb *give* or, as has been sometimes suggested, restricted to themes rendered by the third person singular neuter form *it*. These are important insights arising from a corpus-based study. Of equal importance with respect to issues of methodology, however, are the limitations of this study. Due to the very small number of ditransitive clauses with pronominal themes and recipients in the corpora the investigation did not reveal the full range of verbs which may occur in each of the three ditransitive patterns. Nor did it reflect the distribution of the patterns in passive as opposed to active clauses or in all the possible tense/aspect/mood combinations. In fact given the rarity of ditransitive clauses with pronominal themes and recipients suggested by the corpora we considered, it is not at all clear how large a corpus would need to be to feature all the distributional data which would allow one to fully characterize each of the three constructions. In short, while a corpus is an excellent methodological

---

[2] For more information about the BNC see e.g. Aston and Burnard (1998).
[3] For more information see Siewierska and Hollmann (2005).
[4] See http://www.anglistik.unifreiburg.de/ institut/lskortmann/FRED/ (12 Octoiber 2005)/
[5] See http://www.eng.helsinki.fi/ varieng/team3/1_3_4_2_hd.htm (12 October 2005).
[6] See http://www.gmcro.co.uk/other/NWSA/nwsa.htm (12 October 2005); cf. also Siewierska and Hollmann (2005).

tool for identifying the presence of certain types of constructions in a language or dialect and gauging some idea of their distribution, it cannot be used as the sole source of data in the case of infrequently occurring patterns, at least not if a detailed characterization of the constructions in question is what is being sought. This need to move beyond corpus data becomes even more apparent in relation to the analysis of two other variables of Lancashire dialect to which we now turn.[7]

4. Moving beyond corpora: past tense BE, the definite article, and the need for modern elicitation task data

Our Lancashire data exhibit a considerable amount of variation, both inter-speaker and intra-speaker. Whereas the former can to some extent be successfully accounted for by carefully designed corpora, the latter constitutes much more of a challenge for corpus-oriented methodologies, as we will now proceed to show.

Past tense BE in Lancashire is sometimes realised in the standard way, but sometimes there is levelling of the paradigm. We have observed levelling towards *was* (see examples (2a-b), below), which is quite common in other British English dialects as well (cf. Tagliamonte 1998 or Britain 2002 for a convenient summary), but also towards *were* (see examples (3a-c), below), which is less common in other varieties, especially in non-negative polarity contexts.

(2a)   *…I'd have have a workman then that couldn't understand what you* was *saying but this other Pakistani had learned him* (NWSA, FB)

(2b)   *If the ships* was *coming in he used to be watching them come in and watching 'em going out.* (NWSA, TC)

(3a)   *It's a long time since I* were *up.* (NWSA, ED)

(3b)   *… and they thought he* were *going to die but he fooled them…* (NWSA, HF)

(3c)   *Now Billy* were *a part time rat-catcher.* (NWSA, JA)

Interestingly, while the data we have studied so far is too restricted to draw any confident (i.e. statistically significant) conclusions, it would appear that the use of *were* throughout the paradigm is actually more frequent than that of *was*. Of the 18 speakers in our corpus, 9 display levelling in past tense BE with some degree of regularity, i.e. between 20% and 70% of the time (the remaining speakers only use nonstandard forms in 3% of the cases or less). Of these 9 speakers, 6 level to WERE

---

[7] Another respect in which our study of ditransitives has been limited so far is that we have not considered other factors. Yet issues such as information structure may play a role as well in the ordering of theme and recipient. A more sophisticated follow-up study could take this into account by using multifactorial analysis (see e.g. Gries 2005 for an example). The variables discussed in Section 4 may also involve additional factors apart from dialect (we are thinking here especially of the token frequency of the construction in question), which again future work may take into consideration.

more often than to WAS, 1 speaker levels to WERE and WAS with more or less equal frequency, while only 2 level to WAS more often than to WERE.

Variation is also observed in the use of the definite article. There are three variants: (i) the article is realised fully, as in standard English, (ii) it is reduced (usually realised as a glottal stop, normally represented orthographically as *t'*), or (iii) it is omitted altogether. These deviations from the standard occur in some other Northern dialects as well, e.g. Yorkshire (see e.g. Barry 1972; Jones 2002, 2004). Using recordings from the Survey of English Dialects Jones (2004) suggests that reduction to zero is restricted to a small area in the North East, around Kingston Upon Hull.[8] Our NWSA data, however, show that it is not uncommon in Lancashire, though it is too early to say whether this is true for all or only part of the county.

One central question concerning the distribution of any linguistic variable is its degree of 'sociolinguistic salience'. In the following we aim to show that this degree can be determined on the basis of evidence from elicitation protocols. As an indicator of sociolinguistic salience we use the susceptibility of a variable to trigger accommodation. A method to quantify accommodation is proposed that is based on the distribution of standard vs. non-standard variants relative to the duration of an interview.

Variation, social values and accommodation

When we see a speaker oscillating between standard and non-standard forms, it is natural to hypothesise that the alternation may be due to accommodation.[9] Specifically, the hypothesis is that as the interview progresses the speaker approximates their normal, casual speech style and therefore avoids the non-standard variant less and less. It is important to note that this hypothesis rests on the assumption that the variable is subject to some degree of more or less conscious control. In Labov's (1972, 1994, 2001) terms, if we see stylistic variation then the variable is a so-called *marker*. If, on the other hand, distribution of the variants is not stylistically stratified, then the variable is a mere *indicator*. In the latter case there may still be social variation, but speakers are utterly unaware of it. Labov's hierarchy of sociolinguistic sali-

---

8 See also http://www.yorksj.ac.uk/dialect/DARmorvar.htm (12 October 2005), map 5.

9 The standard assumption in the literature is that speakers accommodate to their interlocutors (see e.g. Bell 1984; Giles, Mulac, Bradac, and Johnson 1987; Street and Giles 1982; Thakerar, Giles, and Cheshire 1982). Accommodation is a complex phenomenon, encompassing not only accent/dialect but also, for instance, speed and loudness. In terms of accent/dialect, however, Paul Kerswill (p.c.) argues that the reality may be more complicated than the traditional assumption. He points out that it may be more accurate to say that speakers accommodate to their perception of the interlocutor's variety. Since in our NWSA data the interviewer's variety is usually closer to standard English than that of the interviewee, and because of the inherently formal character of an interview situation, the interviewees may be adjusting their speech to standard English or to their perception of the standard. For our purposes it is not necessary to take a position on what exactly is the target of the accommodation - the crucial observation is that the general direction is away from casual speech.

ence includes a third value, i.e. *stereotype*. A variable is stereotyped if speakers are so aware of it that they comment on it overtly, and correct or even hypercorrect it. Diachronically, a variable's degree of salience need not remain constant: in discussing the difference between changes from above and changes from below, Labov suggests that in the former the level of social awareness is high from the outset. Changes from below, on the other hand, are characterised by a gradual climb up the hierarchy of salience indicator<marker<stereotype (see e.g. Labov 2001, 196-7).

Various researchers have commented on the question as to what factors determine the degree of sociolinguistic salience associated with a given linguistic variable. Kerswill and Williams (2002) offer a useful overview of much of the literature. They point out that in order for salience to help explain language variation and change, it must be grounded in some independent factors. Some of these factors are social, while some others are broadly perceptual-cognitive, e.g. frequency (cf. Bardovi-Harlig 1987), a transparent relation between form and meaning (Mufwene 1991, 139; Chapman 1995, 2-3; also known as *isomorphism* in the iconicity literature, cf. e.g. Haiman 1980), prosodic salience (see e.g. Yaeger-Dror 1993) and interactive prominence (operationalised by Cheshire 1996:6 as encompassing interrogative and negative contexts). There is an obvious analogy between size in the visual field and phonetic substance in language processing, so this seems another possible determinant of salience. While concurring with Kerswill and Williams on the relevance of social factors, we see these cognitive-perceptual factors as primary. The reason is that linguistic items will normally be more or less free from social values when they come into existence. It is only after they have emerged that social forces can start working on them (cf. also Croft's 2000 strict distinction between 'innovation' and 'propagation', where the role of socio-historical linguistics is restricted to the latter stage of a given change). In other words, we suggest that ultimately it is the cognitive-perceptual constraints that make a form more or less liable to becoming subject to social evaluation and patterning.

So far scholarly work has focused on individual perceptual-cognitive determinants; the question as to whether and how they work together has not been addressed. This is because the various determinants have been treated as applying on a specific level of linguistic organisation only: Yaeger-Dror (1993) invokes prosodic salience in accounting for allophonic variation, Mufwene (1991) and Chapman (1995) use isomorphism in accounting for morphological processes, and Bardovi-Harlig (1987) and Cheshire (1996) argue that frequency and interactive prominence, respectively, play a role in variation between grammatical constructions. By contrast, we take the position that in accounting for syntactic variation there is no a priori reason why for instance isomorphism could not play a role. Our research question, then, is whether it is possible to shed light on the relative importance of this set of constraints in grammatical variation.

Our Lancashire data allow us to operationalise this question: by comparing processes of accommodation across various variables — with different perceptual-cognitive properties such as token frequency, form-meaning relation, and phonetic

substance — we may be able to tease out which factors contribute more to salience in grammar, and which ones are less important.

Measuring accommodation

The operationalisation process requires a method to quantify accommodation. Our solution is to first determine the position of each variant in the whole text, and then calculate whether the use of the non-standard variant as against the standard variant increases towards the end of the text. Concretely, each variant token receives a score that is equal to the number of words of preceding text. The scores for the standard and non-standard uses are then compared with the Mann-Whitney U-test (one-tailed; see e.g. Butler 1985:98-102).[10] (If you could just insert a reference to one of these statistic textbooks. I think that that is what he is after. And if someone doesn't know the test at all, they can look it up.)If non-standard variants turn up significantly more often towards the end of the text, we conclude that in the beginning the interviewee adjusted their speech to the (perceived) standard.[11]

So far we have restricted our analysis to the past tense of BE and to the use of the definite article. The data for ditransitives with pronominal objects is as yet too sparse to enable any conclusions to be drawn. An additional consideration in leaving out these constructions is the potential inapplicability of the notion *variable* to syntactic phenomena because of the lack of full semantic equivalence between variants, a position entertained in the sociolinguistic literature (see e.g. Kerswill and Williams 2002, 100). We feel that this may be too strong, and would prefer to see each case studied on its own terms. With respect to ditransitives, however, we would agree that the 'variants' are quite possibly not exactly synonymous (e.g. in terms of topicality). This problem does not apply to the past tense of BE and to the use of the definite article. As for the latter, another reason for looking at this variable is Jones's suggestion that "[t]he most stereotypical feature of northern British English dialects, especially those of Yorkshire and Lancashire [is] the occurrence of vowel-less forms of the definite article" (2002, 325). It is unclear whether Jones intends his use of the word "stereotypical" in the Labovian sense (see above), but at any rate we might expect to obtain some evidence for accommodation.

As we noted in Section 3 we have at this stage only a few corrected NWSA interviews available, and for the variables under consideration a high degree of accuracy is crucial: consider that the difference between, for example, a full definite article token and a reduced one is only slight. Of the corrected scripts we have at our

---

10 The U-test is not a very powerful test. It can even take interval variables as its input — while position in the text as defined in our calculation can be seen as a ratio variable. This means that if we detect a statistically significant difference, we can be very confident indeed that this is a meaningful result.
11 This method is a clear example of by-subjects analysis, which Gries (this volume) argues - rightly, in our view - is not separated from by-items analysis often enough in corpus linguistics.

disposal only two speakers who use the non-standard forms of these variables with any regularity. No firm conclusions can therefore be reached as yet, but that is unproblematic given the methodological focus of this paper.

The frequencies of standard and non-standard variants of the two variables under consideration in the two speakers are given in Table 1, below. For past tense BE we distinguish between levelling to *was* (represented here as '*was* Pl') and levelling to *were* ('*were* Sg').

		TC[12]	ED
Past tense BE	*was*	81	10
	was Pl	16	2
	were	12	77
	were Sg	11	95
Definite article	full	145	47
	reduced	33	24
	zero	15	52

Table 1: Standard vs. non-standard forms of past tense BE and the definite article in two Lancashire speakers

In our analysis of BE we made a further distinction — not shown in the table — between existential and non-existential uses. Another parameter that cannot be read off the table is, of course, the positions of all the tokens in the texts. Our statistical analysis of the distribution of these tokens shows that in neither of the two speakers is there any evidence for accommodation in relation to the definite article: the standard vs. non-standard tokens are evenly spread.[13] Nor do the data for speaker TC display a skewed distribution of the use of standard vs. non-standard forms of past tense BE. In ED, by contrast, there is a very highly significant (i.e. $p<.001$; U=369) difference in her use of *were* in singular existentials (as in example (4) below) compared to the standard use of *was* in this context. As the interview progresses, the proportion of non-standard uses increases. This accommodation effect is still significant if we pool all uses of past tense BE in ED, although only at the lower level of $p.<.05$ (U=4988).

(4)   *There* were *like a kind of a bit of a nursery where that aquarium is now 'cos I was surprised when I saw that.* (NWSA, ED)

Should the pattern observed in ED turn out to be robust across more transcripts, then at least for Lancashire that would raise problems for Jones's suggestion that the

―――――――――

12 The speakers are identified by their initials. By participating in the interviews carried out by the North West Sound Archive, they have consented to their information being made publicly available.

13 We have compared the full form against the reduced and zero realisations together, against both of these two separately, and have also compared the reduced form against the omitted article.

definite article is the most stereotyped feature. And in relation to the determinants of salience mentioned above, it would also mean that token frequency alone is not enough to render a variable salient: in the spoken part of the BNC there are around 120,000 instances of past tense BE, less than a third of the definite article (c. 410,000 tokens).

Let us briefly illustrate how a more extensive study may shed light on the other factors, again assuming that the accommodation effect in past tense BE proves to be robust but not in the case of the definite article. In terms of prosodic salience, *was/were* clearly rank higher than the definite article. *The* only receives heavy emphatic stress under exceptional circumstances (e.g. where there is some sort of contrast with an earlier noun phrase), and can neither appear at the end of a tone unit, nor form one on its own. Therefore, the prosodic salience of past tense BE may well help explain why it should be subject to social evaluation more easily than the definite article. Phonetic substance may play a role as well. While it is not straightforward to measure *was/were* vs. *the/t'/Ø* precisely and unambiguously, one may say that past tense BE can, and often is, realised as three phones whereas the definite article never amounts to more than two. Moving on to isomorphism, the importance of this factor is not easy to assess, as it depends on how many meanings one wishes to assign to past tense BE and to *the*. If one were to take a very coarse-grained view, and argue that the auxiliary and copula uses of BE are two meanings, whereas the definite article is always associated with 'givenness', then the accommodation facts would run counter to what one might expect on the basis of isomorphism. But this is not the only possible semantic analysis of these variables. Interactive prominence, finally, is also problematic. Controversies arise already in relation to Cheshire's definition of the factor itself, which is sensitive to interrogatives and negatives but not, for example, imperatives — despite the fact that all three do exhibit an association with different speech acts. However, even if we accept this definition, we run up against problems, especially if we rely on corpus data. In the two transcripts we analysed there are only 24 instances of either past tense BE or the definite article used in interrogative or negative contexts. Such low numbers make statistical analysis impossible. The problem here is largely inherent in the context in which the data were obtained: interviewees as opposed to interviewers only very rarely ask questions. And since the task is to relate experiences from the past, there is not much need for negating propositions either. In order to investigate the importance of interactive prominence, then, we will need to elicit — along the lines suggested by the SAND researchers (see e.g. Cornips and Jongenburger 2001, Cornips and Poletto 2005) — questions and negatives (as well as perhaps imperatives) — a corpus alone will not do.

5. Conclusion

Before we end our brief discussion of the different methodologies required in the investigation of grammatical variables let us briefly consider the issue of the object of our study itself, i.e. the regional variety, or varieties, of Lancashire. Purely linguistic definitions of dialects have long been recognised as highly problematic (Chambers and Trudgill 1998, Chapter 1). It seems to make more sense to think of dialects as social constructs, as is done especially in perceptual dialectology (e.g. Preston 1989, 1998). By getting informants to draw dialect areas Preston has been able to arrive at social definitions of dialects in the US. He uses a computational method to work out the correspondences between people's perceptions. This allows him to come up with a "generalised mental dialect map" (e.g. Preston 1989, 344). We envisage a similar approach for our project. Whilst a purely linguistic definition of Lancashire is doomed to fail, a generalised mental map will represent a socially adequate definition. Such a map should then have implications for the localities included in the further stages of our study: we would be inclined to select sampling points not on the basis of the administrative county boundary, but the psychological one.

To conclude, we hope to have shown that corpora have been a great asset in dialectology, and have in fact been key in the emergence of the lively field of dialect syntax. While we will continue to use them extensively, we recognise at the same time that we will benefit from modern elicitation tasks (and interpretation methods). Finally, in order to be able to define our object of study properly we will embrace the third recent trend in studies on regional variation, i.e. perceptual dialectology.

Works Cited

Aston, G. and L. Burnard (1998). *The BNC Handbook: Exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.

Bard, E.G., D. Robretson and A. Sorace (1996). "Magnitude Estimation of Linguistic Acceptability." *Language* 72:32-68.

Bardovi-Harlig, Kathleen (1987). "Markedness and Salience in Second-Language Acquisition." *Language Learning* 37:385-407.

Barry, M. V. (1972). "The Morphemic Distribution of the Definite Article in Contemporary Regional English." M. F. Wakelin, ed. *Patterns in the Folk Speech of the British Isles*. London: The Athlone Press.

Beal, Joan and Karen Corrigan (2002). "Relativisation in Tyneside English." P. Poussa, ed. *Relativisation on the North Sea Littoral*. Munich: Lincom Europa, 33-56.

Beal, Joan and Karen Corrigan. (2005). ""No, Nay, Never", Negation in Tyneside English." Y. Iyeiri, ed. *Aspects of English Negation*. Amsterdam: John Benjamins, 139-56.

Bell, A. (1984). "Language Style as Audience Design." *Language in Society* 13, 145-204.

Bock, J.K. (1986). "Syntactic Persistence in Language Production." *Cognitive Psychology* 18, 335-387.

Britain, David (2002). "Dialect Contact and Past Be in the English Fens." Ms., Department of Language and Linguistics, Essex University.

Bucheli, Claudia and Elvira Glaser (2002). "The Syntactic Atlas of Swiss German Dialects: Empirical and Methodological Problems." Sjef Barbiers, Leonie Cornips, Susanne van der Kleij, eds. *Syntactic Microvariation*. Amsterdam: Meertens Intitute, 41-73. <http://www.meertens.knaw.nl/ books/synmic/> (12 October 2005).

Butler, Christopher S. 1985. *Statistics in Linguistics*. Oxford: Basil Blackwell.

Butler, Christopher S. (2003). *Structure and Function. A Guide to Three Major Structural-Functional Theories.* Amsterdam/Philadelphia: John Benjamins.

Chambers, J.K. and P. Trudgill (1998). *Dialectology*. Cambridge: Cambridge University Press.

Chapman, Carol (1995). "Perceptual Salience and Analogical Change: Evidence from Vowel Lengthening in Modern Swiss German Dialects." *Journal of Linguistics* 32, 1-13.

Cheshire, Jenny (1996). "Syntactic Variation and the Concept of Prominence." Juhani Klemola, Merja Kytö and Matti Rissanen, eds. *Speech Past and Present: Studies in English Dialectology in Memory of Ossi Ihalainen*. Frankfurt: Peter Lang, 1-17.

Cornips, Leonie and Willy Jongenburger (2001). "Elicitation Techniques in a Dutch Syntactic Dialect Atlas Project." Ton van der Wouden and Hans Brokehuis, eds., *Linguistics in the Netherlands 2001*. Amsterdam/Philadelphia: John Benjamins, 53-63.

Cornips, Leonie and Cecilia Poletto (2005). "On Standardizing Syntactic Elicitation Techniques (part 1)." *Lingua* 115, 939-957.

Cowart, Wayne (1997). *Experimental Syntax: Applying Objective Methods to Sentence Judgments*. Thousand Oaks: Sage Publications.

Croft, William (2000). *Explaining Language Change.* Harlow: Longman.

Gerritsen, M., ed. (1991). *Atlas van de Nederlandse Dialectsyntaxis* (AND). Amsterdam: P.J. Meertens-Instituut voor Dialectologie, Volkskunde en Naamkunde.

Giles, H., Mulac, A., Bradac, J. J., and Johnson, P. (1987). "Speech Accommodation Theory: The First Decade and Beyond." M. L. McLaughlin, ed., *Communication Yearbook 10*. Beverly Hills: Sage, 13-48.

Greenbaum, Sydney (1973). "Informant Elicitation of Data on Syntactic Variation." *Lingua* 31, 201-12.

Gries, Stefan Th. 2005. "Syntactic Priming: A Corpus-Based Approach." *Journal of Psycholinguistic Research* 34, 365-99.

Haiman, John (1980). "The Iconicity of Grammar: Isomorphism and Motivation." *Language* 56, 515-40.

Ihalainen, Ossi. (1991). "On Grammatical Diffusion in Somerset Folk Speech." Peter Trudgill and J.K. Chambers, eds. *Dialects of English. Studies in Grammatical Variation*. Harlow: Longman, 104-19.

Jones, Mark (2002). "The Origin of Definite Article Reduction in Northern English Dialects: Evidence from Dialect Allomorphy." *English Language and Linguistics* 6, 325-45.

Jones, Mark (2004). The Phonetics and Phonology of Definite Article Reduction in Northern English Dialects. PhD diss., Cambridge: University of Cambridge.

Kerswill, Paul and Ann Williams (2002). ""Salience" as an Explanatory Factor in Language Change: Evidence From Dialect Levelling in Urban England." Mari C. Jones and Edith Esch (eds.), *Language Change. The Interplay of Internal, External and Extra-Linguistic Factors*. Berlin/New York: Mouton de Gruyter, 81-110.

Kortmann, Bernd, Tanja Herrmann, Lukas Pietsch, and Susanne Wagner, eds. (2005). *A Comparative Grammar of British English Dialects. Agreement, Gender, Relative Clauses*. Berlin/New York: Mouton de Gruyter.

Labov, William (1972). *Language in the Inner City: Studies in the Black English Vernacular.* Philadelphia: University of Pennsylvania Press.

Labov, William (1994). *Principles of Linguistic Change. Volume 1: Internal Factors.* Oxford: Blackwell.

Labov, William (2001). *Principles of Linguistic Change. Volume 2: Social Factors.* Oxford: Blackwell.

Mufwene, Salikoko (1991). "Pidgins, Creoles, Typology, and Markedness." Francis Byrne and Thom Huebner, eds. *Development and Structures of Creole Languages. Essays in Honor of Derek Bickerton*. Amsterdam: John Benjamins, 123-43.

Orton, H., M.V. Barry, W.J. Halliday, P.M. Tilling and M.F. Wakelin (1963-69). *Survey of English Dialects*. 4 vols. Leeds: E.J. Arnold.

Preston, Dennis R. (1989). *Perceptual Dialectology: Nonlinguists' Views of Areal Linguistics.* Dordrecht: Foris.

Preston, Dennis R., ed. (1999). *Handbook of Perceptual Dialectology, Vol. I*. Amsterdam: John Benjamins.

Schütze, Carson T. (1996). *The Empirical Base of Linguistics. Grammaticality Judgments and Linguistic Methodology*. Chicago: University of Chicago Press.

Shorrocks, Graham (1999). *Grammar of the Dialect of the Bolton Area. Part II. Morphology and Syntax.* Frankfurt: Peter Lang.

Siewierksa, Anna and Willem Hollmann (2005). "Ditransitive Clauses in English with Special Reference to Lancashire Dialect." Mike Hannay and Gerard Steen eds. *The English Clause: Usage and Structure*. Amsterdam/Philadelphia: John Benjamins.

Street, R.L., Jr., and Giles, H. (1982). "Speech Accommodation Theory: A Social Cognitive Approach to Language and Speech Behavior." M.E. Roloff and C.R. Berger, eds. *Social Cognition and Communication*. Beverly Hills: Sage Publications, 193-226.

Tagliamonte, Sali (1998). "*Was/Were* Variation across the Generations: View from the City of York." *Language Variation and Change* 10, 153-91.

Tagliamonte, Sali. and H. Lawrence (2000). ""I Used To Dance, But I Don't Dance Now." The Habitual Past in English." *Journal of English Linguistics* 28, 323-53.

Thakerar, J. N., Giles, H., and Cheshire, J. (1982). "Psychological and Linguistic Parameters of Speech Accommodation Theory." C. Fraser and K. R. Scherer, eds.

Advances in the Social Psychology of Language. New York: Cambridge University Press.

Wagner, Susanne (2005). "Gender in English Pronouns: Southwest England." Bernd Kortmann, Tanja Herrmann, Lukas Pietsch and Susanne Wagner, eds. *A Comparative Grammar of British English Dialects. Agreement, Gender, Relative Clauses*. Berlin/New York: Mouton de Gruyter, 211-367.

Yaeger-Dror, Malcah (1993). "Linguistic Analysis of Dialect "Correction" and Its Interaction with Cognitive Salience." *Language Variation and Change* 5, 189-224.