

Syntactic variation and spoken languageⁱ

Jenny Cheshire

Queen Mary, University of London

Mile End Road, London E1 4NS

UK

J.L.Cheshire@qmul.ac.uk

1. Introduction

Linguists who analyse spoken language have often commented on the challenge that it poses for conventional analyses of syntactic structure. An early comment of this kind was from Crystal (1976:166), whose analysis of clause structure in spontaneous conversational English led him to claim that the linguistic organization of this variety of English had been “fundamentally misconceived”. Crystal attributed our lack of understanding partly to the absence of data and partly to the uncritical application of traditional paradigms of enquiry. We can no longer claim an absence of data, now that electronic corpora of transcribed speech are readily available; but it is still not always recognised that we cannot easily understand the nature of spontaneous spoken language in terms of the standard linguistic descriptors (Milroy 2001:270). In this paper I discuss two fundamental characteristics of spoken language that do not fit well with traditional paradigms and that in my opinion are relevant to a potential alliance between the fields of variationist linguistics and generative linguistics.

The first characteristic is the high proportion of prefabricated expressions that occur in spontaneous unplanned speech. This is relevant to generative approaches to syntactic variation because what may appear to be syntactic structures to be explained within the framework of the internal grammar may be better seen as prefabricated expressions with little flexibility. They would be better accounted for, therefore, within another component of the grammar, or perhaps at an interface level. Their relevance for or variationist linguistics is that what we may wish to analyse as a syntactic variant may be

more appropriately seen as a lexical item. I will discuss some examples in section 4.

The second characteristic of spoken language to be considered here is the pervasiveness of affective meanings. These can be difficult to accommodate within conventional generative and variationist frameworks alike, but they need to be taken into account since the expression of affective meanings may influence the extent to which variation arises in syntax. I will give some examples in section 5. The examples in sections 4 and 5 will also illustrate some of the difficulties that arise if different syntactic structures are treated in variationist research as surface realisations of one and the same variable.

I will argue that a discourse-oriented analysis is a necessary complement to the analysis of syntactic variation, firstly because it can help to identify structures that are prefabricated and secondly because it can show how speakers use syntactic structures in spoken interaction. This in turn may give some insights into when, and why, variation occurs. I begin however by noting the lack of attention to the structure of spoken language within both the generative and the variationist traditions.

2. The generativist approach and spoken language

Generative theory aims to characterise our innate knowledge of language structure. It is assumed that this knowledge is identical for the individual speakers of a language (indeed, at a more general level, for speakers of all languages), and that appropriate data can come from a single speaker's intuitive judgements about the grammaticality of constructions. The theory has not been concerned with performed, externalised language, so it has not been

necessary to consider either the syntactic organisation of spoken language or the potential differences between spoken and written varieties of language. However, now that generativists have begun to work with data arising from variationist research it is impossible not to consider these questions, for the data used in the study of variation and change come from what in this tradition is considered externalised language. Variationists prioritise spoken rather than written language (Schneider 2002: 67), and there has been a tradition of analysing wherever possible the most spontaneous unplanned ‘vernacular’ speech, where speakers pay the minimum attention to its production and the Observer’s Paradox is least in evidence (Labov 1970). This for ‘internal’ and ‘external’ approaches to be mutually supportive rather than conflicting, as Chomsky has recently argued that they should be (Chomsky 1999: 34), it is necessary to consider the extent to which the generative model of syntax is appropriate for the analysis of spoken language as well as for the analysis of data obtained from elicited intuitions.

There are arguments in favour of seeing the generative model as appropriate for externalised language as well as for data obtained from intuitions. For example, the intuitions of speakers of languages with a strong written tradition, such as English, are likely to be heavily influenced by the written language (for example, Brown and Yule 1983). Furthermore, some corpus linguists have argued that there is a shared common core between the syntax of spoken and written language (see, for example, Leech 2000). This would suggest that data obtained from intuitions may fit well with the structure of externalised language, both written and spoken. It can also be claimed that generative theory has developed in part from the intuitions of speakers of

languages that do not have a written form. If we can accept that speakers have an intuitive knowledge of the structures they produce when speaking, this is a further reason for claiming that the model is applicable to spoken language is strengthened.

However, several researchers who have analysed corpora of spoken language claim that the structures of spoken language differ both from data obtained from intuitions and from the syntax of planned written language. For example, spoken English – but not planned written English – contains utterances such as (1) and (2) below, uttered within a single intonation contour (see Cheshire 1999). Here two clauses appeared to be fused: in (1) these are *that's really what Professor Galbraith was talking about* and *what Professor Galbraith was talking about is that there's a huge knock-on effect*:

(1) that's really what Professor Galbraith was talking about is that
there's a huge knock-on effect

(2) It's just a bunch of baloney is what it is.

Crystal (1976) gives more complex examples, including some involving intercalation, as in (3). The clauses are set out on separate lines as in Crystal (1976:158) for ease of description.

(3) a. I'm very suspicious of the press generally

b. and I can tell you

c. because not only I mean that's one case

d. that you've given

- e. but also on their reporting of erm affairs foreign affairs
- f. because living in Cyprus
- g. I've seen quite a number of historical events you know

Here the clauses in (3c–e) are subordinate to the clause in (3a), and those in (3f–g) are subordinate to the clause in (3b). Crystal points out that from a semantic viewpoint (3e) relates to both (3a) and (3f): it is half of the reason for the proposition expressed in (3a), along with (3c) and (3d), but it also constitutes a new theme linking (3c) with (3f) and (3g). The syntactic status of (3e) is unclear: there is no main verb and it could, Crystal argues (*op. cit.*: 158), be analysed as a complex adverbial linked via the because of (3f) to (3g).

There are several accounts now of syntactic forms that are specific to spoken language. These include, for English, Biber et al. (1994), Miller (*in press*); for English and Russian, Miller and Weinert (1998); and for French, Blanche-Benveniste (1997). It is not clear how structures such as these could be considered as generated by the grammar, yet many researchers working outside the generative paradigm have argued that they should be included within a model of grammar. Schegloff (1989: 143), for example, argues that a biological perspective on language should examine language in the natural environment in which it occurs (in other words, in unplanned conversation), and should entertain the idea that the syntactic structures of spoken language are adaptations to this environment. A growing body of work shows that spoken language has its form partly as a result of interactional factors, such as the turn-taking mechanism (see, for example, Ochs et al. 1996, Couper-Kuhlen and Selting 2001, and Ford 1996), and that these factors help determine variation.

For example, Ford (op. cit. 147-148) finds that variation in the clause position of conditional and temporal adverbial clauses in American English is related to their discourse functions and the conversational context. They are more frequently in clause-initial position in extended spans of talk, where speakers have negotiated a special right to the floor and where there are more opportunities for the discourse-structuring functions of the clauses to be realised. Mondorf (2000) adds a sociolinguistic perspective: in the London-Lund corpus of spoken British English male speakers tend to position adverbial clauses initially, whereas female speakers place them more frequently in clause-final position. Mondorf explains these preferences in terms of the different orientations of male and female speakers to information management and epistemic grounding. Levinson (1983: 97) has argued more generally that social principles for co-operative interaction, such as those shown in politeness strategies, have a pervasive effect on language structure.

It has been claimed, then, from several different quarters, that interactional and social factors can constrain both the form of spoken syntax and aspects of syntactic variation. It is not yet clear to what extent the structure of spoken syntax can be explained as the result of performance mechanisms that do not need to be accounted for within the internal grammar (as it is conceived by generativists), but it becomes difficult to avoid the question when data from spoken language are used to develop generative theory.

3. The variationist approach and spoken language

Since researchers working in the variationist tradition have always worked with the data of 'externalised' language, they might be expected to

have paid attention to the characteristics of spoken syntax. However here too the nature of spoken language has been largely neglected, albeit for different reasons.

One reason is that the linguistic variable was originally conceived for the analysis of phonological variation. Variants should be semantically equivalent: in other words, they should be alternative ways of 'saying the same thing' (Chambers and Trudgill 1998: 50). Semantic equivalence can be easily established for phonological variables, where the form-meaning relationship is at its most arbitrary, but there has been much controversy about whether it can also be established for syntactic variation. The issues were much discussed during the 1970s and 1980s (see, for example, Lavandera 1978, Cheshire 1987, Levinson 1988, Romaine 1980, Weiner and Labov 1983), and debate has continued since then (see, with reference to French, Blanche-Benveniste 1997, Coveney 1997, Gadet 1997; and for general discussions Cheshire, Kerswill and Williams in press, Cornips and Corrigan in press, Coveney 2002, Milroy and Gordon 2003). A tacit consensus seems to be that the condition of strict semantic equivalence can be relaxed for syntactic variables, so that a variable can be set up on the basis of an equivalence in discourse function (Dines 1980, Coupland 1983). For example, the five forms in (4) to (8) below (from Romaine 1984: 426) can all be considered ways of 'saying the same thing' in that they all have the same communicative intent or discourse function (the speaker would like the addressee to close the window):

(4) it's cold in here

(5) I'm cold

(6) are you cold?

(7) would you close the window?

(8) close the window

The problem however is that we are now dealing with variation that is constrained by pragmatic factors rather than by the grammar. It is not clear to what extent forms that we might wish to consider as examples of syntactic variation are always motivated by pragmatic factors, and there is no consensus in the field on whether this invalidates their analysis as linguistic variables (see Winford 1996:188 for further discussion). The issues are no longer much discussed within the variationist literature, but the legacy of the debate accounts in part for the neglect of syntactic variation relative to phonological variation.

A second reason for the neglect of spoken syntax in variationist research is that the methodology does not require a detailed analysis of syntax.

Researchers focus on one linguistic variable at a time, extracting tokens from the conversational contexts in which they occur. For example, an analysis of *was/were* variation in an English dialect involves extracting and coding all the tokens of *was* and *were* produced by the speakers participating in the study, and then performing a quantitative analysis to determine the linguistic and social constraints on the occurrence of *was* rather than *were*, and vice-versa. Once the variable to be analysed has been selected, there is no need to consider the syntax other than to identify which potential internal constraints are to be included in the statistical analysis. Different researchers have chosen for themselves what to include as a potential constraint, and decisions are not always based on systematic syntactic grounds (see Henry 2002). Analysts working on *was/were* variation, for example, usually take account of agreement

between the subject and the *was/were* form, and the polarity of the construction in which the forms occur (in some dialects negation favours *weren't* rather than *wasn't*), but the focus of the analysis remains the *was/were* variable and the procedure does not necessarily require a fuller analysis of the syntactic construction in which *was/were* occurs, nor of the overall discourse structure. Syntactic constructions specific to spoken language, therefore, such as (1) to (3) above, have tended to be overlooked, as have the possible effect of interactional factors.

A related reason for the neglect of spoken syntax in this field is that variationists tend to analyse the same grammatical variables over and over again. This is partly because the favourite variables tend to occur with the high frequencies that are necessary for quantitative analyses. However another, less obvious, reason seems to be that analysts have been influenced by the ideology of the standard. *We was*, for instance, is non-standard relative to standard English *we were*, and this has made the form salient not only to prescriptivists and laypeople but also to linguists. In addition, the sociocultural processes involved in standardisation means that the use of non-standard forms correlates with a speaker's social status. Variables of this type are eminently suitable for analysis within the variationist framework since they meet the classic definition of a sociolinguistic variable (a structural unit with two or more variants involved in co-variation with social variables). Variationists have worked almost exclusively on languages that have been heavily standardised, so the potential influence of the standard ideology on the selection of variables for analysis has been high (see Cheshire and Stein 1997 and Milroy 1999 for further discussion). Thus for English the most frequently analysed variables are

morphosyntactic forms where one variant is prescriptively non-standard: they involve subject-verb agreement, negative concord and non-standard negative forms such as *ain't*, as well as various standard and non-standard verb forms. Syntactic variation involving less salient forms have been largely ignored.

Both generativists and variationists, then, have largely neglected the structure of spoken language. Despite the different methodologies that they traditionally adopt – with generativists relying on intuitions and variationists on audio recordings of spoken interaction – each approach has analysed forms abstracted from the interactional context in which they occur. A further similarity is that each approach focuses on abstract linguistic systems, with generativists aiming to characterise the structures of the internal grammar and variationists seeking to understand the properties of variable systems (Milroy and Gordon 2003: 8). An essential difference, of course, is that generativists assume that the grammar is identical for all speakers of a language. Their concern, therefore, is to identify which aspects of the observed variation are categorical, and how this variation can be accommodated within the grammar. Variationists on the other hand are also concerned with understanding the social embedding of variation, especially insofar as this can explain the processes by which orderly linguistic change occurs. Variation for them is integral to the nature of language and must be incorporated within the model. In principle, therefore, variationists are equally interested in social (external) and internal constraints on variation, though different studies may tend to give more weight to one than the other².

In any event, for both variationists and generativists the analysis of syntactic variation rests on an appropriate identification of the forms that are

considered to alternate. I hope to show in what follows that a discourse-oriented approach can help to identify these forms, and that this is a necessary complement to generative and variationist analyses.

4. Prefabricated expressions

Prefabricated expressions in spoken language

Many researchers working on spoken language have claimed that linguists tend to over-emphasise the creative aspect of language. There is no doubt that we can produce and understand an infinite number of sentences that we have never heard before but, as Bolinger (1975: 297) pointed out, the fact that we *can* do this does not mean that we do. It would be counter-productive in spontaneous face-to-face communication to constantly produce brand new sentences, and speakers use prefabricated expressions to help them cope with the demands of fast speech production. These expressions include conversational routines with clear social or cultural functions, such as forms conventionally expressing apologies, thanks, compliments or requests (for example the English request formulae *I wonder if I could/ could you possibly/ can I just*), frequent collocations, like *heavy smoker, white coffee*, ‘construction templates’ such as *as far as I (can see/know/can make out)*, or sentence builders such as *my point is, I’m a great believer in* (see Crystal 1995: 162ff, Wray 2002).

Estimates of the proportion of ready-made chunks of unanalysed language in large-scale corpora of spoken language range from 30 per cent (Biber et al. 1999) to 70-90 per cent (see Aijmer 1996:31). The difference in the estimated proportions reflects the ways in which the chunks are defined.

Sometimes researchers rely on subjective identifications of what counts as prefabricated, whilst others give a strict definition on the basis of collocation patterns within in a large corpus. For example, Biber et al.'s (1999) analysis of a 40 million word corpus of spoken and written English excludes combinations of less than three words: it therefore excludes recurrent noun and adjective combinations such as *heavy smoker* and recurrent conversational routines like *I'm sorry*. Estimates of the proportion of prefabricated expressions also reflect decisions about how fixed in form an expression must be in order to be considered prefabricated. *How do you do*, for example, is completely frozen; and the 'sentence builder' (Crystal 1995: 162) *what I mean is* is capable only of limited alteration (such as *what I really mean is*, *what I meant to say was*).

Some conversational routines have greater flexibility; these include, for example, the compliment formula *I (really) like/love your NP*, where the NP must refer to an item that is culturally approved (Holmes 1995). Aijmer (1996: 217) accounts for the flexibility of certain conversational routines by seeing them as 'mini-grammars' consisting of collocational stems generating a limited set of structures. An example is the expression *to put it another way*: in the London-Lund corpus this could be described as having a stem generating the related discourse forms *putting it*, *put it*, and *put*, followed by one of four manner adverbials (*this way*, *like this*, *another way* and *mildly*). The interrogative *how shall I put it* also occurred in the corpus. Aijmer proposes that conversational routines can be arranged along a continuum from completely fixed forms through semi-fixed forms (e.g. *I'm so/really/very sorry*), frame and slot forms (e.g. *could I have X*) to mini-grammars. It is

difficult to accommodate mini-grammars within a formal grammar, however, because their output is so constrained.

Even a strict definition of what constitutes a prefabricated expression gives their proportion within a corpus as 30 per cent: a sufficiently high proportion for their existence to be taken seriously. They raise the question of whether spoken language might be better conceptualised as linear and sequential in structure rather than as hierarchical. The idea is pushed to its limits by Sinclair (1991: 68), who predicts that “lexical hordes will invade the traditional domain of syntax “ and lead to its eventual demise Skehan (1998: 37) takes a more moderate view, suggesting that the production of speech involves improvising on a clause by clause basis, such that speakers use lexical phrases and lexical sentence stems wherever possible in order to minimize processing demands, and only as a last resort generate language that is not part of our memorised lexicon. Even a moderate view, however, suggests that when we are analysing spontaneous spoken language it is important to bear in mind that what may appear to be a syntactic construction may instead be a chunk of ready-made memorised language. I will consider some examples from spoken English in the following section.

Some prefabricated expressions in spoken English

Independent adverbial phrases in spoken English are a case in point. These appear to be subordinate adverbial clauses in that they are introduced by conjunctions such as *because*, *when*, or *if*, but there is no main clause. Generative theory does not allow for the possibility of unattached adverbial clauses – understandably, since by definition an adverbial clause is subordinate to a main clause (and they may well be overlooked by researchers, since it is

not clear that constructions such as these are accessible to our intuitions). However, Both Mondorf (2000) and Ford (1993) noted unattached adverbial clauses in their analyses of adverbial clauses in spoken English. Mondorf recorded 6 per cent out of the total number of adverbial clauses (259, out of 4462 clauses); and Ford found 3 per cent out of the total number of temporal adverbial clauses (2 out of 63 temporal clauses). Both authors were able to infer a main clause from the surrounding linguistic context, but it is not always possible to do so. McCarthy (1998:79-82) for example notes clauses introduced by *if* and *cos* (a reduced form of *because*) that occur alone and function as main clauses. I found it equally impossible to infer a main clause for some *when* structures that occurred in a corpus of conversations between 12–16-year-old working-class adolescents in Reading, Berkshire (see Cheshire 1982). Unattached phrases introduced by *when* were relatively frequent in my Reading corpus, accounting for 25 per cent (28) of the 105 *when* clauses. I will discuss these phrases in some detail in order to illustrate the problems they can pose for a variationist analysis.

Two examples of the *when* phrases are indicated by the arrows in (9) and (10). They were uttered with level tones on every syllable except the last: this has a falling tone and is slightly drawled. Interestingly, they were used only by the male adolescents.

(9) (*the boys are talking about one of their teachers, who was married to someone I knew. Jenny was the fieldworker/ (me))*)

Nobby: yeah Miss Threadgold she ain't bad

Rob: yeah she . she went camping with us

Jenny: yes he told me she'd been camping

→ Nobby: when we went camping

Rob: she's a good laugh

Jenny: is she?

Nobby: yeah

(10) (*the discussion has been about jobs the girls might consider doing when they leave school*)

Jenny: you have to do horrible jobs if you're a nurse .. all the bed
pans

All: <LAUGHTER>

Jenny: have you ever been in hospital?

Valerie: [I have

Christine: [oh yeah I have

Valerie: I got run over by a car

Christine: I fell off a gate backwards <LAUGHS> and I was
unconscious

→ Tommy: oi when I .. when I went in hospital just for a little while ...

Valerie: sshh

Tommy: cos my sister and my cousin they bent my arm .. they twisted
it right round

A variationist analysis of the *when* phrases would seem in principle to be possible, if we assume one variant to be a *when* clause that is clearly subordinate to a main clause (for example, *when we went camping* in *when we*

went camping we had a great time) and another variant to be an unattached *when* clause as in (9) or (10).

The first step in a variationist analysis would be to establish the discourse function of the lone *when* phrases and the conventional, subordinate, *when* clauses, to ensure that they are equivalent in function. One function of conventional initial *when* clauses is explicatory (Ford (1993: 29, 32). Ford found that this was the case when *when* clauses followed a semantically broad term such as *thing* or *then*. In her data the explication occurred within an extended speaker turn; she argued, in fact, that the use of the semantically broad term contributed to the projection of an extended turn. Only four of the lone *when* clauses in the Reading data were explicatory, however. One of these is illustrated in (11): here Rob explains, in answer to a question, how Britt (one of the playground leaders) tries to control her mind. The lone *when* clause does not elaborate a semantically broad term, nor does it project an extended turn, but it does provide a time frame for a specific situation that illustrates Britt's behaviour. In doing so it clarifies a semantically problematic concept (the idea of controlling your mind) that the emerging discourse has shown to be ambiguous or too vague for present purposes: this was initially unclear to all the participants, as indicated by Rob's *whatever that means* and Nobby's response (*I don't know*) to my question about how this can be done.

(11) Rob: and Britt she's queer = = she's trying to learn to control
her mind

Nobby: = yeah =

Rob: whatever that means

Jenny: is she?

Rob: [yeah

Nobby: [yeah

Jenny: oh how is she going to what is she doing to con

Nobby: I don't know

→ Rob: when you look at smoke and that you know fire =

Jenny: = yeah

Nobby: she looks at a flame she's . you can look at . she's trying to

look at a flame until it burns right out

Jenny: and then w . how does that control your mind?

Rob: I don't know

The four lone *when* phrases with an explicatory function, then, do share at least one of the functions of subordinate *when* clauses.

A further function of subordinate adverbial clauses in initial position is to project an extended turn and present background for material that follows. These characteristics contribute to Ford's view that initial adverbial clauses are pivotal points in the development of talk (op cit: 62). The remaining 21 lone *when* phrases in the Reading corpus share these characteristics. In (10), for example, the other speakers interpret Tommy's lone *when* phrase, prefaced by his attention-getting *oi*, as an indication that he intends to take a projected turn; this is shown by Valerie compliantly telling her younger sister to be quiet. Usually, the extended turns are narratives of personal experience; thus, in (10) Tommy went on to tell the story of his stay in hospital. Both explicatory and pivotal lone *when* phrases, then, share some aspects of the interactional

function of conventional adverbial *when* clauses. As mentioned above, there is social variation in that the forms without an accompanying main clause are used only by the male adolescents. These forms might seem, then, to be candidates for a variationist analysis, with a sociolinguistic variable consisting of two variants, one a *when* clause with a main clause, the other a lone *when* construction without a co-occurring main clause.

However, this approach would miss an important discourse function of the 21 lone *when* constructions that are pivotal in the development of talk. In every case, the narrative that follows the lone *when* phrase concerns events that are familiar to the other speakers, either because they have heard the story before, or because they participated themselves in the events that are recounted. The narrative is a form of joint reminiscing – a discourse event with an important role in reinforcing group membership (Edwards and Middleton 1986). In the Reading playground conversations these narratives were especially significant in the construction and reinforcement of group friendship patterns amongst the male adolescents. The main function of these lone *when* phrases, in other words, is as a story opener, marking the upcoming story as a shared reminiscence. Female adolescents constructed friendships on a more individual basis, telling stories mainly as monologues. Their different narrative style was reflected in their preferred story opener which, as Table 1 shows, was a temporal subordinate clause, clearly situating the story in the past (for further details see Cheshire 2000).

INSERT TABLE 1 HERE

When the lone *when* phrases are considered in their full interactional context, then, it becomes clear that they cannot be analysed as variants of conventional initial *when* clauses, since they are not functionally equivalent. They have a specific discourse function as a story opener marking a shared reminiscence. A variationist analysis could, perhaps, be performed on the range of story openers that are used to introduce sequences of joint reminiscing (those shown as the first group in Table 1), but this of course would hold no interest for the study of syntactic variation.

An analysis that fits better with the data is to see the lone *when* phrases as conversational routines, together with the other story openers marking an upcoming shared reminiscence (such as *what about when, you know when* or *remember when*). As we saw earlier, a conversational routine is a sequence of words that appears to have syntactic structure but that is produced and processed as a more or less prefabricated phrase (Aijmer 1996). The *when* of the lone *when* phrases may be a reduced form of the other *when* phrases in this group of story openers marking shared reminiscences³.

The lone *when* phrases used as story openers are not, of course, completely fixed in their form: they differ, therefore, from prefabricated phrases such as *how do you do?* and are more productive than the *to put it* expressions mentioned in the previous section. Yet they have more in common with prefabricated lexicalised forms such as these than with completely new clauses that have been generated by the grammar. They consist of a frame (*when* + NP + VP, with the verb in the past tense) with a fixed intonation contour. The past tense form of the verb distinguishes the story openers from the other, less

frequent lone *when* phrases with an explicatory function: in (11), for example, the verb *look* is in the present tense. The words that constitute the NP and the VP are repeated from the preceding discourse, and this facilitates their function as a way of taking the floor: thus in (9) Nobby's *went camping* echoes the words of the preceding three turns, and in (10) Tommy's *in hospital* echoes the question *have you have ever been in hospital?*

What might initially appear to be an instance of syntactic variation, then, is more appropriately seen as a conversational routine with an interactional function in turn-taking and a social function in indexing group solidarity (as we have seen, it is used only by the boys, along with other story openers that mark an upcoming shared reminiscence). It is not entirely fixed in form, and conforms more to a phrase generated by a 'mini-grammar'.

Other forms used as story openers in the Reading conversations are better analysed as prefabricated expressions than as constructions generated by the grammar. One such form involves verbal *-s*. This of course is usually considered to be an agreement marker in generative analyses of English, and in present-day standard English it does indeed appear to have this function, occurring only on present tense verb forms with third singular subjects. In many present-day non-standard varieties however the distribution of verbal *-s* differs. In Norwich, England, for example, it is variably absent with third person subjects (Trudgill 1974) ; in Reading it is variably present with non-third person subjects and quasi-categorical with third person subjects (Cheshire 1982), as in several other varieties of British and US English. It is sometimes assumed that speakers have regularised the present tense paradigm in these vernaculars, so that verbal *-s* is an agreement marker in these vernaculars also,

but many researchers have identified a wider, diverse range of functions for the form, perhaps especially in African American English. The functions include marking durative aspect (Pitts 1986, Brewer 1986), habitual aspect (Pitts 1986), variably marking the present tense (Schneider 1983) and marking the historic present (Myhill and Harris 1986): overviews of research on verbal *-s* in English dialects are given by Clarke (1997) and Godfrey and Tagliamonte (1999). Most of these studies exclude from the envelope of variation story openers or topic introducers such as *you know* in (12) and (13). In the Reading corpus, as elsewhere, *you know* used in this way is invariable, never taking the *-s* suffix, unlike *you know* as a lexical verb, as in (14) and (15):

(12) you know that hill down there? I rode down that with no hands on
the

handlebars

(13) you know your mum....you know that bike she had

(14) you knows him don't you Nod?

(15) he says to me "look here and I see if I knows you"

The story opener then, is a prefabricated expression, like the discourse marker *you know* (which performs a range of conversational functions, including adding liveliness to a conversation, and constructing solidarity; see, for example, Holmes 1986, Fox Tree and Schrock 2002: 729. It is not certain, however, that all cases of prefabricated expressions have been accounted for in analyses of verbal *-s*. After all, existential constructions and canonical clause constructions are usually analysed side-by-side, despite their different syntactic

derivations (see Corrigan 1997 for discussion). Analyses of verbal –s, whether generativist or variationist, might benefit from a prior discourse analysis aiming to identify all the prefabricated expressions in which verbal –s occurs: this would not only make the analyses more accountable to the data but would also further our understanding of how and why prefabricated expressions develop and their role in grammaticalisation and other kinds of language change.

5. Affective meanings in spoken language

The prevalence of emotive expressions in conversation is well known (see for example Biber et al. 1999:958), as is the fact that for speakers the communication and construction of affective meaning is as important as the communication of referential meaning. Here I will simply consider the relevance of this fundamental characteristic of language for determining the extent to which syntactic variation may arise in discourse.

As the main example I will consider the pronoun tags that occur in several northern varieties of British English. They are illustrated in (16) and (17) below, indicated by the arrows, with the tag underlined. The extracts are taken from an analysis of variation in the speech of 14-15 year old working-class and middle-class adolescents in Hull, England (Cheshire, Kerswill and Williams 1999). In our data these pronoun tags are used only by the working class adolescents.

(16)a. Charlie: the only time I drink is like at parties or =

b. Matt: = yeah..

not one of the things you do every day really is

it..daft

- c. Charlie: don't like smoking or anything like that ..no that's disgusting
- d. Matt: I used to me..well I tried it
- e. Charlie: I haven't even tried it me
- f. Matt: my mam wouldn't say nowt
- g. AW: do your parents smoke?
- h. Charlie: my mam does
- i. Matt: all of them do..got my real dad my step dad and my mam
- j. Charlie: I don't like it me
-
- (17) a. AW: right what about a favourite singer then?
- b. Kay: Peter André me
- c. Ruth: Peter André's alright but
- d. Kay: he's got a real nice chest him
- e. AW: has he? Is it hairy?
- f. Kay: no it's real brown and greasy
- g. Ruth: cos he has baby oil smothered on him

In (16d), (16e), (16j) and (17d) the tags are co-referential with the subject pronoun in the preceding clause. As such they are subject to aspects of binding theory. Like other discourse-related phenomena considered as dislocation (for example, NP-fronting, *it*-clefts and left dislocation), they can be accommodated in generative syntax within the left periphery of the clause (Henry 1995:135).

Their occurrence in spoken English in Ayrshire has been analysed by

Macaulay (1989, 1991), who found them to be a feature of working class speech. Macaulay reports that the working class speakers in his study used constructions that brought personal pronouns into prominence in a way not found in the speech of the middle classes. These constructions included not only pronoun tags but also left dislocation, NP- fronting, and *it*-clefts, all of which were used more frequently by the lower class speakers with the function of expressing intensity. The middle class speakers, by contrast, tended to convey intensity through the use of adverbials. This previously unsuspected finding has interesting sociolinguistic implications, but it shows the difficulty of using the linguistic variable for the analysis of forms expressing intensity (and Macaulay did not attempt to do so). The variable may well be a heuristic construct that does not necessarily map directly onto the units of linguistic structure (Wolfram 1993) but to include left dislocation, say, or *it*-clefts in the same analytic unit as adverbials would be stretching the concept of the variable beyond all credibility, even if the forms can be considered to have the same discourse function.

Even if we ignore the sociolinguistic differences in the expression of intensity, and focus simply on the tag constructions, it is still difficult to analyse the tags within a variationist framework. Intensity (which seems to be used in the literature with much the same meaning as ‘emphasis’, is often said to be the discourse function of a form that appears to be involved in syntactic variation. Emphasis does not affect truth conditions, so if the pronoun tags in our data are emphatic in function they could in principle be analysed within a variationist framework: the clauses with tags could be seen as semantically equivalent variants to the corresponding clauses without tags. However, it would be

difficult to identify the envelope of variation. Emphasis is an ill-defined concept, lacking theoretical rigour. It does not provide a basis for predicting which variant will be preferred on any one occasion, nor even where it is possible for variation to occur (Schwenter in press); and to say that emphasis highlights a particular entity in the discourse – in this case, the co-referential subject pronoun – is too subjective to be useful if we are looking for language universals, as Myhill (1992:3) points out. Sells et al. (1996:174) claim that unless the status of emphasis can be clearly specified in the grammar, along with the extent to which it can affect the form and function of different linguistic phenomena, the very ubiquity of appeals to this type of affective meaning may reduce its analytic value.

Nevertheless speakers do appear to choose a range of expressions to add some kind of additional, affective meaning, to their propositions. We are omitting an aspect of language that is important to speakers if we do not take the expression of emphasis into consideration when attempting to explain syntactic variation, whether within a generative or a variationist framework.

A further factor constraining a speaker's use of pronoun tags is their interactional function. This too, however, it can be difficult to take into account in a rigorous way. For example, in the Hull data the tags sometimes occur with a form referring to an entity that becomes a conversational topic: in (17b), for example, Kay's *Peter André* is picked up by Ruth in the next turn. Similarly, in (17e) Ann Williams picks up the topic of Peter André's nice chest, proposed by Kay in the preceding turn. In both these examples, then, the tag has a function in conversational management. Note that in (17d) the fact that the tag is co-referential with the subject pronoun is irrelevant in terms of its interactional

function: it is the entire proposition (Peter André and his chest) that is proposed as a topic.

Elsewhere in the data the tags sometimes appear to explicitly mark a contrast, often between the content of the utterance in which they occur and the content of the previous turn. This is possibly the case in (16d), where Matt and Charlie are discussing smoking with Ann Williams. Charlie is a keen anti-smoker, and his first utterance about smoking, in (16c), makes it clear that he does not like it (*don't like smoking or anything like that...no that's disgusting...*). Matt, with his *I used to me* in (16d) perhaps shows that he has inferred from this that Charlie does not smoke – or perhaps he already knows this to be true – and he claims, in contrast, to have smoked himself. Charlie's *I haven't even tried it me*, in (16e), then clearly contrasts his own lack of experience with Matt's, and in (16j) he repeats the assertion he made in (16c), this time contrasting his own dislike of smoking with the behaviour of Matt's family, which has been described in the preceding two turns.

The tags may simultaneously (or alternatively) signal a contrast between what the speaker would find it interesting to talk about and what the previous speaker has been saying: thus from (16c) onwards both Matt's and Charlie's contributions concern their own experiences of smoking and their family's behaviour concerning smoking. The tags may signal a desired change of topic, then, often in contrast with a previous topic, but there is no principled way of deciding on the basis for the contrast. In Charlie's turns he also expresses his stance towards the content of his clause; the tags may draw attention to his stance and contrast it with Matt's. There are many interactional factors, then, that can affect a speaker's choice of construction and that may constrain

variation between clauses with tags and clauses without tags.

Many other constructions are similarly conditioned by interactional factors. A further example from English is the *get*-passive construction, illustrated in (18):

(18) Josephine got run over by a bus

This can be seen as an alternant to a conventional *be*-passive, as in (19), and both (18) and (19) as alternants to the corresponding active clauses in (20) and (21):

(19) Josephine was run over by a bus

(20) A bus ran Josephine over

(21) A bus ran over Josephine

Weiner and Labov (1983:43) claimed that a shift to the *get*-passive is one of the most active grammatical changes taking place in present-day English. An important question for the study of language variation and change therefore concerns the distribution of the form throughout the speech community, since this will allow us to chart the diffusion of the change. As with pronoun tags and lone *when* phrases, there is no initial problem in using the variable to analyse this alternation: the alternants in (18-21) have the same verb and the same entities as arguments, so we can assume that they are semantically equivalent.

Passive constructions are relatively infrequent in speech. It is not surprising, therefore, that there have been few quantitative studies of the English passive. Macaulay (1991), however, analysed passive constructions in middle class and working class speech in his Ayrshire study. Although there were no significant social class differences in the overall use of passive clauses, *get*-passives were used more frequently by the working class speakers. *Get*-passives have been a major shibboleth in British schools so it is possible that they were simply avoided by the middle class speakers in the context of the sociolinguistic interview: the working class speakers in the study had far less exposure to formal schooling. Interestingly, however, Macaulay further reports that the *get*-passive occurred almost exclusively with animate subjects and that these, in turn, were also more frequent in the working class interviews. *Get*-passives are eventive aspectually, and this presumably contributes to the animacy effect; events are usually controlled by an actor, and animates are more likely to be able to control events. One factor affecting the use of the *get*-passive, then, is, quite simply, what speakers choose to talk about. If they talk about animate beings, there is a greater chance of their using the *get*-passive.

Carter and McCarthy (1999) add a further dimension to our understanding of the use of this form: their corpus-based analysis reveals that the *get*-passive highlights the stance of speakers towards the grammatical subject and the event encoded in the verb phrase – a stance that usually indicates their judgement that the circumstances are adverse, problematic or otherwise noteworthy. We cannot necessarily assume that the Ayrshire speakers and the speakers in Carter and McCarthy's study are affected equally by these factors, but the two studies at least suggest the possibility that the shift

to the *get*-passive in present-day English is led by a group-specific discourse preference for talking about animate beings and encoding syntactically the speaker's stance towards actors and the event referred to by the verb. Is there, in other words, a distinctive habitual pattern of interaction for the working class group of speakers that is not shared by the middle class group? Unfortunately, however, we now run into the same problems as with appeals to emphasis: although a wide range of syntactic features have been said to express the speaker's stance towards the content of their utterances (see, for example, those listed in Ochs and Schieffelin 1989) it is difficult to give a rigorous definition of the concept and it has no status within any syntactic theory of language.

Syntactic variation and syntactic change seem here to be intimately and inextricably part of the social construction of discourse. Carter and McCarthy's findings point to a similar conclusion: as these authors say (op.cit: 55), judgements about adversity, noteworthiness and the like are socio-culturally founded and emergent in the interaction rather than inherent in the semantics of verb choice or the selection of voice or aspect. In order to address these issues, then, and to fully analyse patterns of variation and change in the use of English passive constructions, it is necessary to add a qualitative, interpretive dimension to the analysis.

All syntactic variables in fact, including low level variables such as verbal *-s*, call for this kind of interpretive discourse analysis. Levinson (1988:166) made this point in relation to English *ain't*, asking whether working class speakers who use *ain't* frequently do so because for them *ain't* is a marker of group identity, or because it is a more emphatic form of negation than *isn't*, *aren't*, *hasn't* and *haven't*. If the latter, does this reflect the habitual patterns of

social interaction of the social group to which the heavy *ain't* users belong? They might, for example, utter emphatic denials more frequently than other social groups in the community because they more often receive accusations. Thus in order to understand how and why speakers use variation, and the effect that their usage has on language change, we cannot simply analyse the simple alternation of forms: we must also perform qualitative analyses to see how these forms are used in social interaction and find a way of incorporating this dimension into an explanatory theory of language structure and language change.

6. Conclusion

My starting point in this paper was that the syntactic structure of spoken language has not been seriously taken into account by either variationists or generativists. A hope, from my sociolinguist corner, is that a new alliance between variationists and generativists might lead to generative linguists becoming more interested in externalised, performed language – not simply in order to find new data against which to test the theory but also to apply the rigour of the generative approach to discovering the structure of spoken language. There are many ways in which our understanding of spoken language might benefit. First, it would be helpful if some theoretical rigour could be brought to the concepts of emphasis (or intensity) and stance, which seem so important to speakers and which are so often appealed to in descriptions of language use. Second, as pointed out by Henry (2002: 277) generativists might help variationists determine on a more systematic basis those syntactic structures that should be considered variants of a single form.

This would help assess the universal dimension of any internal constraints on variation (Cornips and Corrigan in press) and may prevent us from being unduly influenced by the standard ideology when deciding on the forms to be analysed as variants of a single variable. It might even help resolve the issue once and for all of whether the variable should be used to analyse variation in syntax. Third, it may be possible to identify some language universals governing the form of spoken language – though for this to be possible we would need, in my view, to broaden our perspective and look for social or interactional principles in addition to the principles governing innate structures. Perhaps in this endeavour we will need to work with researchers from other fields of linguistics. For example, the perspective of language has been very successfully applied to the findings of social dialectology (see, for example, Kortman 2002, 2004)⁴

Generativists and variationists might work together to develop a methodology that could determine when the phrases of spoken language have been generated by the grammar and when they are prefabricated sequences produced from memory. A decision needs to be taken about whether constructions that appear to be semi-flexible and that are capable of description within ‘mini-grammars’ are best accounted for as generated by the internal grammar or as part of the lexicon (or, as suggested earlier, at an interface level). A related problem is that of establishing, for any given syntactic feature that seems to be variable, how much of the variation is most usefully considered as syntactic (in other words, internally motivated by the grammar) and how much is better seen as pragmatically determined.

Many of these issues will require both a qualitative and a quantitative dimension to the analysis. I have argued in this paper for a complementary approach to the analysis of syntactic variation, which analyses syntactic forms in their discourse context. If we are to gain insights from our different research traditions we need to be aware that the way t speakers use the forms of spoken language may result as much from interactional and social factors as from biological factors and come to a principled decision about whether such forms should be seen as generated by the grammar or attributed to performance mechanisms. In this way we may at last succeed in understanding how the cognitive and the social aspects of language are integrated as part of the human experience.

Note

1. I would like to warmly thank David Adger, Miriam Meyerhoff, Lesley Milroy, Leonie Cornips and Karen Corrigan for comments on this paper, as well as the anonymous referees. All shortcomings are my own responsibility.
2. In fact analyses of syntactic variation have tended to give more weight to internal constraints than social constraints, perhaps because of the problems inherent in using the variable to analyse syntactic variation; see Milroy and Gordon (2003).
3. In all these expressions *when* appears to have indefinite meaning (Lyons 1977, Haspelmath 1997); it resembles its use in conversation as a deictic that is non-specific for one of the participants (as in, for example, *tell me about the*

time when you were lost in the mountains, meaning ‘the time you were lost’, or
I tell you what .. I’ll wash the dishes if you do the cooking.

4. The typological approach is of course divergent from and perhaps even runs counter to the aims of the present volume.

References

- Aijmer, Karin. 1996. *Conversational Routines in English: Convention and Creativity*. Harlow: Longman.
- Biber, Douglas. Conrad, Susan. and Reppen, Randi 1994. “Corpus-based approaches to issues in applied linguistics.” *Applied Linguistics* 15, 169-187.
- Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan and Finegan, Edward 1999. *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Ltd.
- Blanche-Benveniste, Claude. 1997. “La notion de variation syntactique dans la langue parlée.” *Langue Française* 115, 19-29.
- Bolinger, Dwight 1975. “Meaning and memory.” *Forum Linguisticum* 1, 2-14.
- Brewer, Jerome P. 1986. “Durative marker or hypercorrection? The case of –s in the WPA ex-slave narratives.” In *Language Variety in the South*, Michael B. Montgomery and Guy Bailey (eds.), 131-148. Tucson: The University of Alabama Press.
- Brown, Gillian. and Yule, George 1983. *Discourse Analysis*. Cambridge: Cambridge University Press.
- Carter, Ronald and McCarthy, Michael 1999. “The English *get*-passive in spoken discourse: description and implications for an interpersonal grammar.” *English Language and Linguistics* 3, 41-58.

- Chambers, J.K. and Trudgill, Peter 1998. *Dialectology*. Cambridge: Cambridge University Press. 2nd edition.
- Cheshire, Jenny 1982. *Variation in an English Dialect: a Sociolinguistic Study*. Cambridge: Cambridge University Press.
- Cheshire, Jenny 1987. "Syntactic variation, the linguistic variable and sociolinguistic theory." *Linguistics* 25, 257-282.
- Cheshire, Jenny 1999. "Spoken standard English." In *Standard English: The Widening Debate*, T. Bex and R. J. Watts (eds.), 129-148. London: Routledge.
- Cheshire, Jenny 2000. "The telling or the tale? Narratives and gender in adolescent friendship networks." *Journal of Sociolinguistics* 4, 234–262.
- Cheshire, Jenny, Kerswill, Paul and Williams, Ann. 1999. *The Role of Adolescents in Dialect Levelling*. Final Report submitted to the Economic and Social Research Council, London, June 1999. Ref. R000236180.
- Cheshire, Jenny, Kerswill, Paul and Williams, Ann in press. "On the non-convergence of syntactic and phonological variation." In *Dialect Convergence and Divergence in a Changing Europe*, Peter Auer, Franz Hinskens and Paul Kerswill (eds.). Cambridge: Cambridge University Press.
- Chomsky, Naom (1999). Derivations by Phase. Unpublished MSS. Boston, MIT.
- Clarke, Sandra 1997. English verbal –s revisited: The evidence from Newfoundland. *American Speech*: 72: 227-259.

- Cornips, Leonie and Corrigan, Karen in press. "Convergence and divergence in grammar." In *Dialect Convergence and Divergence in a Changing Europe*, P. Auer, F. Hinskens and P. Kerswill (eds.), Cambridge: Cambridge University Press.
- Corrigan, Karen 1997. *The Syntax of South Armagh English in its Socio-historical Perspective*. Unpublished doctoral dissertation, University Collge Dublin.
- Couper-Kuhlen, Elizabeth and Selting, Margret (eds.) 2001. *Studies in Interactional Linguistics*. Amsterdam: Benjamins.
- Coupland, Nikolas 1983. "Patterns of encounter management: Further arguments for discourse variables." *Language in Society* 12, 459-476.
- Coveney, Aiden 1997. "L'approche variationniste et la description de la grammaire du français." *Langue Française* 115, 88-100.
- Coveney, Aiden 2002. "Supplement: A critical survey of recent research on grammatical variability, negation and interrogation in French." *Variability in Spoken French*. Bristol: Intellect.
- Crystal, David 1976. "Neglected grammatical factors in conversational English." In *Studies in English Linguistics*, S. Greenbaum, G. Leech and J. Svartvik (eds). 153-166. London: Longman, pp. 153-166.
- Crystal, David 1995. *The Cambridge Encyclopedia of the English Language*. Cambridge: Cambridge University Press.
- Dines, Elizabeth R. 1980. "Variation in discourse – "and stuff like that"". *Language in Society* 9, 13-33.

- Edwards, John and Middleton, Derek 1986. "Joint remembering: constructing an account of shared experience through conversational discourse." *Discourse Processes* 9, 423-459.
- Ford, Cecilia E. 1993. *Grammar in Interaction: Adverbial Clauses in American English Conversations*. Cambridge: Cambridge University Press.
- Fox Tree Jean E. and Schrock, Josef C. 2002. "Basic meanings of *you know* and *I mean*". *Journal of Pragmatics* 34:727-747.
- Gadet, Françoise 1997. "La variation, plus qu'une écume." *Langue Française* 115, 5-18.
- Godfrey, Elizabeth and Tagliamonte, Sali 1999. Another piece for the verbal –s story: Evidence from Devon in south-west England. *Language Variation and Change* 11: 87-121.
- Haspelmath, Martin 1997. *Indefinite Pronouns*. Oxford: Oxford University Press.
- Henry, Alison 1995. *Belfast English and Standard English: Dialect Variation and Parameter Setting*. New York: Oxford University Press.
- Henry, Alison 2002. "Variation and syntactic theory." In *Handbook of Language Variation and Change*, J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), 267-282. Oxford: Blackwell.
- Holmes, Janet 1995. *Women, Men and Language*. Harlow: Longman.
- Holmes, Janet 1986. "Functions of *you know* in women's and men's speech". *Language in Society* 15: 1-21.
- Kortmann, Bendt 2002. "New prospects for the study of dialect syntax: Impetus from syntactic theory and language typology." In *Syntactic*

Microvariation, S. Barbaiers, L. Cornips and S. van der Kleij (eds.).
Amsterdam.

Kortmann, Berndt 2004 (ed.) *Dialectology meets Typology: Dialect Grammar from a Cross-linguistic Perspective*. Berlin/New York: Mouton de Gruyter.

Labov, William 1970. "The study of language in its social context." *Studium Generale* 23: 30-87.

Lavandera, Beatriz 1978. "Where does the sociolinguistic variable stop?" *Language in Society* 7, 171-182.

Leech, Geoffrey 2000. "Grammars of spoken English: The new implications of corpus-oriented research." *Language Learning* 50, 675-724.

Levinson, Stephen 1988. "Conceptual problems in the study of regional and cultural style." In *The Sociolinguistics of Urban Vernaculars*, Norbert Dittmar and Peter Schlobinski (eds.), 161-190. Berlin: Walter de Gruyter.

Levinson, Stephen 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lyons, John 1977. *Semantics. Volume II*. Cambridge: Cambridge University Press.

Macaulay, Ronald K. S. 1989. "He was some man him': Emphatic pronouns in Scottish English." In *Synchronic and Diachronic Approaches to Linguistic Variation and Change*, Thomas J. Walsh (ed.), 179-187. Washington DC.: Georgetown University Press.

Macaulay, Ronald K. S. 1991. *Locating Dialect in Discourse: The Language of Honest Men and Bonnie Lassies in Ayr*. New York: Oxford University Press.

- McCarthy, Michael 1991. *Discourse Analysis for Language Teachers*.
Cambridge: Cambridge University Press.
- McCarthy, Michael 1998. *Spoken Language and Applied Linguistics*.
Cambridge: Cambridge University Press.
- Miller, Jim in press. "Spoken and written English." In *The Handbook of English Linguistics*, B. Aarts and A. McMahon (eds.). Oxford: Blackwell.
- Miller, Jim and Weinert, Regina 1998. *Spontaneous Spoken Language: Syntax and Discourse*. Oxford: Clarendon Press.
- Milroy, Lesley 1999. "Standard English and language ideology in Britain and the United States." In *Standard English: The Widening Debate*, T. Bex and R. J. Watts (eds.) 173-206, London: Routledge.
- Milroy, Lesley 2001. "Conversation, spoken language and social identity." In *Style and Sociolinguistic Variation*, P. Eckert and J.R. Rickford (eds.), 268-278. Cambridge: Cambridge University Press.
- Milroy, Lesley and Gordon, Matthew 2003. *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell.
- Mondorf, Britta 2000. *Sex differences in English syntax*. Unpublished doctoral dissertation, Düsseldorf University
- Montgomery, Michael, Fuller, Janet M. and DeMarse, Sharon 1993. "The black men has wives and Sweet harts [and third pliral -s] Jest like the white men': Evidence for verbal -s from written documents on 19th century American speech." *Language Variation and Change* 5: 335-357.
- Myhill, Jim 1992. *Typological Discourse Analysis*. Oxford: Blackwell.

- Myhill, Jim and Harris, Wendell A. 1986. "The use of the verbal –s inflection in BEV." In *Diversity and Diachrony*, David Sankoff (ed.), 25-32. . Amsterdam/Philadelphia: Benjamins.
- Ochs, Elinor, Schegloff, Emmanuel A. and Thompson, Sandra A. (eds.) 1996. *Interaction and Grammar*. Cambridge: Cambridge University Press.
- Ochs, Elinor and Schieffelin, Bambi 1989. "Language has a heart". *Text* 9, 7-25.
- Pitts, Ann 1986. "Contrastive use of verbal –s in slave narratives." In *Diversity and Diachrony*, David Sankoff (ed.), 73-82. Amsterdam/Philadelphia: Benjamins.
- Poplack, Shana and Tagliamonte, Sali 1989. "There's no tense like the present: Verbal –s inflection in early Black English." *Language Variation and Change* 1: 47-84.
- Quirk, Randolph, Greenbaum, Sydney, Leech, Geoffrey and Svartvik, Jan 1985. *A Grammar of Contemporary English*. Harlow: Longman.
- Romaine, Suzanne 1980. "The relative clause marker in Scots English: Diffusion, complexity and style as dimensions of syntactic change." *Language in Society* 9, 221-247.
- Romaine, Suzanne 1984. "On the problem of syntactic variation and pragmatic meaning in sociolinguistic theory," *Folia Linguistica* 18: 409-439.
- Schegloff, Emmanuel 1989. "Reflections on language, development and the interactional character of talk-in-interaction." In *Interaction in Human Development*, Marc H. Bornstein and Jerome S. Bruner (eds.), 139-153. Hillsdale, N.J.: Lawrence Erlbaum.

- Schneider, Edgar 1983. "The origin of the verbal –s in Black English."
American Speech 58: 99-113.
- Schneider, Edgar. 2002. "Investigating variation and change in written documents." In *Handbook of Language Variation and Change*, J. K. Chambers, Peter Trudgill and Natalie Schilling-Estes (eds.), 67-96. Oxford: Blackwell.
- Schwenter, Scott (in press). "No and *Tampoco*: A pragmatic distinction in Spanish negation." *Journal of Pragmatics*.
- Sells, Peter, Rickford, John, and Wasow, Thomas 1996. "Variation in negative inversion in AAVE: an optimality theoretic approach." In *Sociolinguistic Variation: Data, Theory and Analysis. Selected Papers from NWAVE 23*, Arnold, Jennifer et al. (eds.), 171-1760. Stanford: CSLI.
- Sinclair, John 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Skehan, Peter 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Trudgill, Peter 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Weiner, Judith and Labov, William 1983. "Constraints on the agentless passive." *Journal of Linguistics* 19, 29-58.
- Wolfram, Walt 1993. "Identifying and interpreting variables." In *American Dialect Research*, Dennis R. Preston (ed.), 193-222. Amsterdam/Philadelphia: John Benjamins.

- Winford, Donald 1996. "The problem of syntactic variation." In *Sociolinguistic Variation: Data, Theory and Analysis. Selected Papers from NWAWE* 23, Arnold, J, et al. (eds.) 177-192. Stanford: CSLI.
- Wray, Alison 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Skehan, P. 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Trudgill, P. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Weiner, J. and Labov, W. 1983. "Constraints on the agentless passive." *Journal of Linguistics* 19, 29-58.
- Wolfram, W. 1993. "Identifying and interpreting variables." In *American Dialect Research*, D. R. Preston (ed), 193-222. Amsterdam/Philadelphia: John Benjamins.
- Winford, D. 1996. "The problem of syntactic variation." In *Sociolinguistic Variation: Data, Theory and Analysis. Selected Papers from NWAWE* 23, Arnold, J, et al. (eds.) 177-192. Stanford: CSLI.
- Wray, A. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Table 1. Story openers in the corpus of narratives

<u>Total</u>	<u>All-female conversations</u>	<u>All-male conversations</u>
<u>Markers of a shared reminiscence</u>		
<i>Remember when</i> clause 1	0	1
<i>What about that time when</i> 2	0	2
<i>You know when</i> clause 1	0	1
<i>What about X</i> 1	1	0
Pivotal lone <i>when</i> -clause 5	0	5
<i>I can't forget that time when</i> 1	0	1
Total markers	1	10
<u>11 temporal subordinate clauses introduced by:</u>		
<i>once</i> 6	6	0
<i>when</i> 7	5	2
<i>the other day</i> 5	5	0
<i>one time</i> 2	2	0
<i>one day</i> 3	3	0
<i>last time</i>	1	0
<i>yesterday</i> 1	1	0
Total temp. sub. clauses	23	2
<u>Zero</u> 43	9	34
<u>Miscellaneous</u>		
clause <i>right</i> 8	1	7
left dislocation 12	9	3
<i>there was X</i> 9	6	3
<i>you know X</i> 4	2	2

<i>you see</i>	3	0
3		
<i>see</i>	2	0
2		
<i>you should have seen X mate</i>	0	1
1		
<i>fuck me</i>	0	1
1		
<i>he's a bastard mate</i>	0	1
1		
<i>oh it's horrible</i>	1	0
1		
<i>it wasn't half fun</i>	1	0
1		
<hr/>		
Total all story openers	58	64
122		

Transcription Conventions

..	short pause (not timed)
=	utterance latched on to previous turn (with no discernible pause)
?	question marks show the end of a stretch of talk interpreted as an question
<LAUGHTER>	angled brackets give additional information
[extended square brackets show the beginning of an overlap
→	an arrow indicates that the line to the right is the one where a given example occurs
