NEGLECTED GRAMMATICAL FACTORS IN CONVERSATIONAL ENGLISH
DAVID CRYSTAL

The judicious blend of meticulous observation, experimental enquiry and theoretical insight concerning the English language, which characterizes the writing of Randolph Quirk, has regularly provided linguistic theory with both an example and a challenge. The example lies especially in the complementary roles given to corpus, acceptability test, and intuition in the elucidation of problems. The challenge is to construct a model which will satisfactorily account for the diverse data which these different methodologies uncover. Nowhere is this challenge more disconcerting than in the need to provide an account of that English variety assumed to be the basis of our routine behaviour – informal domestic conversation – for it is here that the discrepancy between standard descriptive statement and observed reality is most noticeable. The aim of the present paper, accordingly, is to identify some of the neglected linguistic features of this variety, which will have to be incorporated into standard descriptions if this discrepancy is to be removed.

The reasons for the comparative neglect of domestic conversational analysis are well-recognized. It is difficult – if not impossible – to capture the informal spontaneity of this variety using a process of controlled intuitive reflection. And corpus-based techniques are problematic, because of the difficulty of obtaining uncontaminated samples of data – uncontaminated either by observer presence, informant self-consciousness, or poor recording quality. The data on which the present paper is based avoided

these problems, using a technique which retained the strengths and eliminated the weakness of the hidden microphone method. Without extremely expensive equipment, this method is of limited value, as the speed and low volume level of much conversation makes for obscure reproduction of one and often all participants. But by using the following strategy, unselfconscious and high-quality recordings were obtained. Friends of the author were invited to his house for a social occasion, but with a specific request to help participate in an 'experiment' on accents. The room was prepared with centrally placed, visible microphones apparently attached to a visible tape recorder; in reality, the microphones were linked to a mixer and recorder in an adjacent room. When the informants arrived, they were given an experimental task to do (such as reciting the alphabet). Once this was completed, the visible tape recorder was ostentatiously switched off, and the microphones pushed back somewhat, but left directly in front of the participants. The hidden recorder was of course permanently on, and thus a good quality record of the relaxed spontaneous speech which subsequently emerged was obtained. Permission to use the material was of course always sought before the tape was replayed, and was always given.

The original aim of the enterprise was to provide advanced learners of English as a foreign language with more realistic samples of conversation than are routinely available in language-teaching courses, and it is in that form that extracts of the material have been published (Crystal and Davy 1975), including the whole of the data used in the present paper.[1] It was in fact the accumulation of unexpected analytic difficulties in processing this material which motivated the present argument. To illustrate the nature of this data, an extract is printed in full below: | marks tone-unit boundary; `'` etc mark direction of nuclear tone; | marks the first pitch-prominent syllable of the tone-unit; ↑ marks a step-up in pitch; ' indicates other stressed syllables (" extra strong stress); · – etc mark degrees of pause length; the word containing the tonic is printed in capitals.

A well |what's the · |what's the 'failure with the ↑FÒOTBALL| I
   mean |this · |this I don't 'really ↑SÈE| I mean it · |cos the
   ↑MÒNEY| · |how 'much does it 'cost to get ÌN| |down the ↑RÒAD|
   |NÒW|

B   I |think it ↑probably – it| 5
   |probably 'is the ↑MÒNEY| for |what you ↑GÈT| you |KNÓW| – erm
   I was |reading in the ↑paper this ↑MÒRNING| a a | CHÀP| he's a
   DI|RÈCTOR| of a |big ↑CÒMPANY| in |BÌRMINGHAM| – who was th
   the |world's ↑number 'one ↑FÒOTBALL 'fan| he |used to ↑SPÈND|
   a|bout a 'thousand a ↑YÈAR| |watching FÒOTBALL| you |KNÓW| 10

1 The tape-collection as a whole is lodged in the files of the Survey of English Usage, University College London. Extracts 1, 3 and 8, the three longest extracts in Crystal and Davy 1975, are used in the present paper (13 minutes of conversation). Reference numbers cite the extract number and line(s) involved.

(C: |CÒO|) – he's he's |watched 'football in ↑every n · on
↑every 'league · 'ground in ÉNGLAND| |all 'ninety TWÓ|
(A *laughs*) – and he's |been to A↑MĒRICA| to |watch ↑West
BRŌMWICH 'playing in A'merica| he's · he's |been to the la
· to |ÒH| · the |LÀST| f f |two or 'three 'world CÙP| · |world   15
CÙP| · mat |THÍNGS| you |KNÓW| · |TÓURNAMENTS| – – and he |goes
to ↑all the 'matches AWÁY| you |KNÓW| |European ↑CÙP 'matches
and 'everything| that |ÈNGLISH teams are PLÁYING in| he's all
'over the ↑WÒRLD 'watching it you SÉE| – |THÌS YÉAR| he's
|watched ↑twenty 'two GÀMES| – |SÒ 'far| |this YÈAR| which is   20
a|bout · FÌFTY per 'cent| of his |NŎRMAL| (C: |good LÒRD|) · and
|even ↑HÈ's getting 'browned ↑ÓFF| and |HÈ was SÁYING| that
erm – you can |go to a NĬGHTCLUB| in |BĬRMINGHAM| – – and
|watch ↑Tony BÉNNET| · for a|bout ↑thirty ↑BŌB| – |something
like THÍS| a |night with ↑Tony ↑BÉNNET| – |have a 'nice ↑MĒAL| 25
· in · |very · ↑plushy SURRŌUNDINGS| very |WĀRM|
|NÍCE| |PLÈASANT| – says it |CÒSTS him| a|bout the ↑SÀME
a'mount of MÓNEY| to |go and ↑sit in a ↑breezy 'windy STÁND| –
(A & C *laugh*) on a · on a |WÒODEN BÉNCH| – to |WÀTCH| a |rather
BÓRING 'game of ↑FÒOTBALL| with |no ↑PERSONÁLITY| and |all   30
DEFÉNSIVE| and |ÈVERYTHING| he |says it's just ↑KÌLLING itself|
you |KNÓW|

Sentences

Any attempt to analyse this data in terms of sentence structure and function
is beset with difficulties from the outset. Sentence identification and
classification is a much greater problem here than in any other variety of
English. Three factors seem to account for the majority of cases:

(*a*) indeterminate connectivity;
(*b*) indeterminate ellipsis;
(*c*) intercalation of structures.

1 Connectivity

The analytic problem is how far a distinction between simple (mono-clausal)
and complex (multi-clausal) sentences can be maintained. In subordinate
clauses, the semantic dependency involved is usually sufficient to guarantee
the obligatoriness of the connective (*eg He came after the man left*), and often
there is a grammatical criterion, regardless of semantics and intonation (*eg
He told us what the answer was*, where the omission of *what* produces an
utterance whose acceptability cannot be salvaged). But regularly in
coordinate clauses, and also in some cases of subordination, there is no clear
semantic distinction between the presence or absence of a connective, and

formal criteria can be found to support the analysis of an utterance as EITHER a single complex sentence OR as a set of simple sentences. The problem is best illustrated by the use of *and*, which introduces nearly one-quarter of all the clauses in the data (see *Table* 1 on *p* 160).[2]

3.115–20

[1] . . . he |gets 'on the 'wrong TRÀIN|

[2] and |ends 'up in the 'wrong PLÀCE| – –

[3] and |finds that he's ↑in a PLÁCE|

[4] that's |perfectly QUĬET|

[5] and |perfectly ÍNNOCENT|

[6] and there's |no ↑STÒRY| –

[7] and |so he 'just ↑WRÌTES 'one| –

[8] and with|in a ↑WĚEK| he's |managed to cre'ate ↑RÌOTS| you |KNÓW|

On syntactic grounds, only the connective in [4] is obligatory. In this variety of English, it is quite normal to have the first part of a clause omitted (*eg looks like another nice day*; *cf* G C E : 545), and this would permit the omission of the *and* in [2], [3] and [5] (*cf* also below). Omitting *and* in [7] and [8] also seems to make no semantic difference, and the grammar and prosody can be used as they are. [6] is somewhat debatable : it might be argued that the *and* here is less likely to be omitted on the grounds that it marks the end of this first sequence of events (='and as a result'); but the intonation, and accompanying linguistic and extralinguistic context, could be used for this purpose instead, and the *and* is certainly not obligatory. Presumably, then, an analysis would pay attention to the optionality of these connectives, and count the above as seven sentences (or eight, if [5] is viewed as clausal, instead of phrasal). The alternative, to call the whole of the above utterance a single complex sentence, is possible, but vacuous (on this basis, some extracts, such as Extract 5 in Crystal and Davy 1975, would have to be considered as containing but one or two sentences, each consisting of several dozen clauses).

If all the cases were as clear as the above, there would be no problem. 3.93–7 illustrates a more difficult case:

[9] they |go to the · 'Ledra 'Palace HOTÈL for EXÁMPLE|

[10] and they |sit at the BĂR| –

[11] and they ab|sorb you know 'one or two FÀCTS| from a |few PÈOPLE|

[12] but |they 'don't 'know the ↑LĀNGUAGE|

[13] and they |don't 'know the ↑PĒOPLE|

[14] and they |don't · ↑really 'know the 'SITU↑ÀTION| –

2 The sample contained 420 clauses, this excluding a further 120 minor sentences and comment clauses (*cf* G C E : 778) and incomplete sentences. All examples are given with clauses on separate lines, each clause being numbered separately.

Here, to make the contrast between [9–11] and [12–14], *but* (or some such phrase) seems essential: omitting *but*, and changing the intonation of [12–14], does not avoid the ambiguity of these lines appearing to be part of the same list as [9–11]. Given the above reasoning, we should therefore conclude that there are five sentences here: [9], [10], [11–12], [13] and [14]. But as [12] is in no way in semantic contrast with [11] alone, but with all of [9–11], this solution is hardly satisfactory.

2 Ellipsis

Unless one wishes to include a general and uncontrollable notion of 'being understood' into one's analysis, it is essential to introduce specific constraints onto the notion of ellipsis. On this basis, it is possible to separate cases such as *Lunch?*, where there is no unique elliptical derivation, and *The man went out and bought a paper*, where there is (*cf* GCE:568, 707 *ff*). In conversational data, however, one frequently encounters cases of an isolated clause or phrase, where it is wholly unclear whether the utterance is colloquially reduced, independent of the linguistic context, or is an utterance in a relationship of ellipsis to some nearby clause. If the latter, it is often unclear which of two competing relationships is correct. An example of all these problems is 1.22–7.

[15] and |HÈ was SÁYING|
[16] that erm – you can |go to a NĬGHTCLUB| in |BĬRMINGHAM| – –
[17] and |watch ↑Tony BÉNNET| · for a|bout ↑thirty ↑BŎB| –
[18] |something like THÍS|
[19] a |night with ↑Tony ↑BÉNNET| –
[20] |have a 'nice ↑MĒAL| · in · |very · ↑plushy SURRŌUNDINGS|
[21] very |WĀRM| |NÍCE| | PLÈASANT| –

In this sequence, several problems arise. Is [21] related to [20] by ellipsis (*which are*), or to [19] (*which is*), or is it a new sentence with colloquial omission of SV (*it is*), or is it an example of 'postponement' (*cf* GCE:963) (*and this is*)? From the point of view of sentence identification and classification, is [21] a separate, coordinate, or subordinate clause? Similarly, is [20] an ellipsis of *you can* (from [16]), *that you can* [16], or even *and you can*? [18–19] are more obviously appositional, to [17]. [15] and [17] have optional *and*, already discussed under **1** above; *that* in [16] is also optional, with the clause following subordinate to [15]. However, the question here is how much subsequent structure is to be analysed as also subordinate to [15]. Are [20–21], with all their problems, also subordinate to [15]? It does not seem possible to choose between these various analyses on empirical grounds. As with **1**, the question of how many and what kind of sentences we are dealing with seems incapable of receiving a definite answer.

3 Intercalation

Frequently in this data, an utterance is produced which seems to contain two interlaced sentences, as in 3.54–60.

[22] I'm |very sus'picious of the PRÉSS| |GÈNERALLY|

[23] and I can |TÈLL you|

[24] be|cause · not 'only I |mean 'that's ÒNE 'case|

[25] that you've |GÌVEN|

[26] but |ÀLSO| |in in their RE↑PÒRTING| of erm af|fairs ↑foreign AF↑FÀIRS| –

[27] be|cause · ↑LÌVING in 'Cyprus|

[28] I've |seen · ↑quite a 'number of HISTÓRICAL E↑VÈNTS| you |KNÓW|

From the context, it is plain that the reason for [22] is given in [24–26]. The reason for [23], *ie* why the speaker is an authority, is given in [27–28]. What we have, therefore, is a structure of the following type:

Main Clause A + Main Clause B + Subordinate Clause A + Subordinate Clause B

and it is this kind of pattern which is here referred to as 'intercalated'. The situation is however more complex than this. [26], from a semantic viewpoint, relates to both sentences: it is half of the reason for [22], along with [24–25], but it also provides the new theme which is the link with [27–28]. Syntactically, [26] has no main verb, and there is thus some motivation for seeing this as a complex adverbial, linked (via the *because* of [27]) to [28]. Because of such complications, we are once again faced with an unclear analysis in terms of sentence structure.

A similar sort of disassociation of structure which raises problems of sentence analysis is 8.48–52:

[29] |I dis'covered that the ↑MǑTHER|

[30] |who had been ↑THÈRE the 'day BEFÓRE|

[31] |wasn't ↑ÌN it|

[32] and |who was ↑now ↑so BĬG|

[33] having |had ↑two LĬTTERS|

[34] that she |couldn't · ↑easily ↑GÈT through the BÁRS| ·

[35] just |wasn't ↑THÈRE|

[36] which was |very ÓDD|

The second relative clause [32] might again be taken as an example of postponement, in view of the main clause conclusion in [31]; but the main clause is repeated in [35], presumably for a mixture of emphasis and clarity of exposition, following the sequence of subordinate clauses [32–34]. We are thus faced with a problem: do we take [35] as a separate sentence, with elided subject, or do we take [31] as an anticipatory performance 'error', or do we recognize a new type of sentence with a 'double' VP?

The problem caused by a lack of correspondence between syntactic and semantic structure is seen again in 3.75–80; the context is an event in Cyprus:

[37] I mean [w] |that's how 'most [pe] 'people ↑TÒOK it|

[38] and erm |so many ↑ÒTHER 'cases| as |WÈLL|

[39] |where there've 'been · erm ↑inter'national 'SITUẄTIONS|

[40] that erm – |people [re] · have ↑really just ↑taken as 'part of their 'normal LĬFE|

[41] and it |"hasn't AF↑"FÈCTED| the |everyday ↑LÌFE of CÝPRUS| at |ÀLL| ·

Here, the sequence sounds like a single sentence; but the initially plausible analysis of [38] as coordinated object of *took* is put in question by the semantic generality (and switch in tense) of [39–40]. Because 'most people' refers to Cypriots, the implication in [38] is that the 'other cases' are also going to be Cyprus-based; but this is unlikely, given the mention of 'international situations' in [39]. It is [41] which takes up the specificity of the Cyprus example (note the use of *it*); though it also shows, inappropriately, the influence of the perfect tense from [39–40]. If this is so, then [38] must be analysed as a new sentence, with elliptical S V, and presumably [41], also, with optional *and*. The overall structure, then, seems to be

Main Clause A + Main Clause B + Subordinate Clause B + Coordinate Clause A

which allows us several options for sentence classification.

These problems are not isolated cases. A more detailed analysis of connectivity items, for instance, shows that of the 420 clauses in the sample, 267 contain at least one connective (64 per cent); there are 322 connectives, and nearly a third of the cases involve 'optional' *and* (see *Table* 1). Moreover, when a distributional analysis is made of the connectives, in terms of whether they occur initially, medially or finally in clause structure, the bias towards initial position is evident: 281 out of 322 connectives are clause initial (87 per cent) (see *Table* 2). If we then exclude the 72 obligatory conjunctions and the 11 cases of optional *that* (which are of a rather different type), we are left with 198 'optional' connectivity features in the sample: in other words, 47 per cent of clauses pose problems of the type discussed above.

It is arguable that all of the above problems arise solely because of the attempt to impose a descriptive model on the data which uses *sentence* as a primitive term. This variety of English, however, does not seem to be readily analysable in terms of sentences. Rather, the *clause* is the unit in terms of which the material is most conveniently organized (as illustrated in the examples above). A model of Clause + connective + Clause . . . makes far fewer assumptions about the organization of the data, and avoids the arbitrariness involved in the discussion of 1–3 above. To work in terms of

Table 1 Frequency of connectives in the sample

OBLIGATORY

Subordinate conjunction	60
Coordinate conjunction	12

OPTIONAL

and	99
but	13
well	18
because	8
Exclamatory (*eg oh*)	12
Comment clauses	
you know	32
I mean	17
Other	11
that, etc (in Object and Relative clauses)	11
Other	29

TOTAL CONNECTIVES	322 in 267 clauses
ZERO CONNECTIVITY	153
TOTAL CLAUSES IN SAMPLE	420

Table 2 Distribution of connectives in clause structure

	Initial	Medial	Final	Zero
Simple sentence	81	9	15	46
Clause in complex sentence	200	9	8	107
Total	281	18	23	153

clauses, moreover, correlates much better with a prosodic analysis of such data, and thus with a possible model of speech production, where the role of intonation (especially the tone-unit) is central (*cf* Laver 1970: 69 *ff*). *Table* 3 shows the correlation between tone-units and clauses in the sample: 54 per cent of clauses are exactly one tone-unit in length (228 out of 420), which is more than twice the frequency of any other correlation.

Adverbials

The lack of clear sentential organization is thus one of the main factors accounting for the discrepancy between conversational data and standard descriptive statement, referred to at the beginning of the paper: almost all such statements insist on the theoretical priority of the sentence. But this is not the only area where the discrepancy is marked. If we look now at the

Table 3 Correlation between clauses and number of tone-units

Number of tone-units	Simple sentence[a]	Clause in complex sentence
less than 1	1[b]	92[c]
1	129	99
1 plus	2	14
2	23	35
3 or more	14	11

[a] Ignoring connectives
[b] This solitary example can be found in Crystal and Davy 1975:1.127.
[c] Largely indirect speech clause sequences

elements of clause structure, a further problem is posed by the category of adverbial, which emerges as a rather more central notion than is to be found in either traditional or most linguistic accounts (an important exception is Jackendoff 1972:47 *ff*; see also Crystal 1966). From a syntactic point of view, adverbials are always considered to be optional elements of clause structure (apart from the few exceptions, such as with *put*, *be*, etc), and unless a special point is being made, they are not usually to be seen in the sentences which constitute the evidence in linguistic papers. The first point to be noted about the present data, then, is the frequency of adverbials: 246 of the 420 clauses contain an adverbial (59 per cent). But the distribution is more interesting (see *Table* 4). Of those clauses NOT containing adverbials, 38 are the introductory clauses of indirect speech (*eg He said X, I think X*, where *X* is usually a clause), and 12 are items of an idiomatic or phatic kind (*eg they're tremendous, that's right, and that was that*). If these are excluded, on the grounds that they do not display the syntactic variation typical of other clause types, the proportion of clauses containing adverbials increases to 66 per cent.

Table 4 Distribution of adverbials in the sample

ADVERBIAL PRESENT		ADVERBIAL ABSENT	
Syntactically obligatory	51		
Semantically obligatory		Indirect speech	38
Main clause	88	Main clause	56
Subordinate clause	31	Subordinate clause	68
Optional modification	71	Phatic	12
Ellipsis	5		
TOTALS	246		174

Table 4 summarizes the main types encountered. Only 51 adverbials are syntactically obligatory, the majority of these co-occurring with the verbs *be*, *go* and *get*, *eg*

[42] and he's |been to A↑MĒRICA| (1.13)
[43] and he |goes to ↑all the 'matches AWÀY| (1.16)
[44] |was it in MADRÌD| (1.54)
[45] |PÀKI-bashing| was – |at its ↑HÈIGHT| |THÉN| I SUP|PÓSE| (3.17)
[46] but he |gets 'on the 'wrong TRÀIN| (3.116)
[47] and |ends 'up in the 'wrong PLÀCE| (3.116)

and an interesting relative clause example

[48] |worst 'game they ↑ever PLÀYED| (1.117)

On the other hand, only 71 of the adverbials are clearly optional, in the sense that their omission would make no difference to the syntactic or semantic acceptability of the clause sequence in which they occur. These can be broadly classified into two types. Firstly, there are adverbials expressing personal emphasis or attitude, *eg*

[49] |this I don't 'really ↑SÈE| (1.2)
[50] |CÓVENTRY 'maybe| (1.67)
[51] |they 'rather 'liked the ↑WÒRD| (3.21)
[52] you |probably DÌD| (3.31)
[53] I'm |very sus'picious of the PRÉSS| |GÈNERALLY| (3.54)
[54] I mean |FÓRTUNATELY| he |WÀSN'T SHÓT| (3.74)

Secondly, there are adverbials which provide detail that is redundant, either because the information is already present elsewhere in the clause, or a previous clause, or it leads nowhere in the subsequent discourse, *eg*

[55] I mean they're pro|gressively ↑getting ↑WÒRSE| (1.39)
[56] I was |reading in the ↑paper this ↑MÒRNING| (1.7)
[57] we'd |taken a 'school 'trip to ÌTALY| (3.14)
[58] and ex|aggerated them ↑out of ↑all PRO↑PÒRTION| (3.49)
[59] I've for|gotten the 'details NÓW| (3.115)
[60] and · |blocked 'up 'one SÍDE| · with |TǏSSUES| (8.67)

It can be reasonably argued, in all such cases, that if the adverbials were not there, no one would have noticed. In isolation, of course, this cannot be justified, but if we restore the context to each of the above sentences, their optionality becomes clear, *eg* in [57] no subsequent reference is made to Italy – it is the school trip which is taken up as the relevant theme.

The remaining adverbial examples in *Table* 4 are all the reverse of this second type. In isolation, they might all be taken as optional, syntactically and semantically, but in context their presence is crucial. Three main types can be recognized:

(*a*) the clause would be ambiguous or false without the adverbial, *eg*

[61] I mean they |never DÒ these 'grounds ÚP| |DÒ they| (1.38)

[62] but there |WÀS an 'interesting 'programme on these 'grounds| (1.52)

[63] |we 'had – to↑matoes in there RÍPENING| (8.104)

(*b*) the meaningfulness of a later clause is dependent solely upon the presence of the adverbial, *eg*

[64] he's a DI|RÈCTOR| of a |big CÒMPANY| in |BÌRMINGHAM| (1.7)
   (*cf* 1.13, 'and he's been to America to watch West Bromwich playing')

[65] |I went to 'Stamford ↑BRÌDGE last year ÓNCE| (1.107)
   (*cf* 1.111 *ff*, where the situation at Stamford Bridge is taken up)

(*c*) the meaningfulness of the clause is dependent on the adverbial recapitulating or contrasting with information from a previous clause, *eg*

[66] did you |GÉT that in 'Cyprus| (3.2)
   (the Cyprus theme had been discussed earlier)

[67] but |people 'went 'on 'living 'quite ↑NŎRMALLY| (3.72)
   (*cf* 3.71, where the point is made about tension in the area)

Sometimes two of these functions can be found in a single adverbial, *eg*

[68] [he said] there was ↑only "one 'modern GRÓUND in ↑ÈNGLAND|
(1.65)

where without *in England* the statement would be unclear, and a subsequent contrast with grounds on the continent would be unmotivated.

Quite often, the item which the first adverbial relates to is itself an adverbial, and an interesting situation of 'mutual dependence' develops, *eg*

[69] [there was] the |sea of – ↑bodies in ↑front of you ↑MÒVING| and |people 'started to PÙSH| BE|HÌND you| (1.125)

[70] |one 'minute there was · 'seventy THÒUSAND in the GRÓUND| and about · |thirty 'seconds LÁTER| or a |minute 'later they were ↑CLÈAR| (1.93)

[71] it was just |boys who went 'round with short ↑HÀIR| |rather 'like · you KNÓW| · |teddy 'boys in the 'mid ↑FÌFTIES| |went 'round with ↑LÒNG 'hair| (3.45)

Semantically obligatory adverbials occurred altogether in 119 clauses (28 per cent), and were particularly common in main clauses (3:1, according to *Table* 4). We may thus conclude that for the data as a whole, 41 per cent of all clauses (170/420) contain an adverbial that is in some sense obligatory. A further 20 per cent of clauses have an adverbial which is optional. The prominence of the adverbial is also underlined by prosody: 78 per cent of all adverbials carry a nuclear tone, and though there are only 241 adverbials in the data, they account for nearly one-third of all the nuclear tones used (of 654 tone-units, 187 have the adverbial carrying the nucleus).

In short, the traditional view of adverbial use, reflected in the standard descriptions of English, seems better reversed: instead of needing a special reason to put an adverbial into a clause, one might say there needs to be a special reason for leaving one out. The clause structure in the present data would be far more satisfactorily accounted for if more attention were paid to the adverbial at an early analytic stage. One might even introduce it obligatorily at such a stage (eg Clause $\longrightarrow$ V + NP + AP), specifying its deletion only in contexts where its use would be incompatible with other features of clause structure, or redundant, in view of the presence of an adverbial in a previous clause. Such an approach might ultimately produce a far more economical syntactic analysis, and a more intuitively acceptable semantic analysis, than one based on repeated application of an optional rule.

Moreover, when one looks in detail at the nature of the NPs in the data, a further contrast with expected descriptive statement emerges: in the sense of 'Premodification + Head + Postmodification', there are only 233 such structures, ie at least 44 per cent of all clauses have no such NP. Table 5 gives the exponence of the pre-verb and post-verb elements of clause structure. The most striking characteristic is the pronoun category: 325 of the clauses have a pronoun or 'empty' word (it, there) as Subject, viz 77 per cent. In post-verbal position, the situation is almost exactly reversed, with 80 per cent of the exponence going to the combination of NP, Adverbial, and Clause (Object/Complement). The end-weight of clauses in English is something which has often been pointed out (eg GCE:943). What has been less remarked for conversational data is

(a) the fact that NPs account for so little clause element exponence (28 per cent), and
(b) the powerful role played by pronouns and adverbials, which together account for 57 per cent of all exponence.

The limited power of statistical reasoning is acknowledged. On the other hand, tendencies such as the above are sufficiently dominant to suggest that

Table 5 Exponence of pre- and post-verbal elements of clause structure

	Pre-verbal	Post-verbal
Pronoun[a]	325	42
NP	43	190
Zero	52	33
Adverbial		113
Clause		32[b]

[a] Including it, there
[b] All the clauses as object of indirect speech verbs

[164]

they cannot be dismissed as 'mere performance'. It is quite possible that an interesting grammar might emerge if a formalization were attempted using the following two rules:

$$\text{Utterance} \longrightarrow \text{Clause} \, (+\text{Connectivity} + \text{Clause})$$
$$\text{Clause} \longrightarrow \text{Pronoun} + V + \begin{Bmatrix} NP+A \\ A \end{Bmatrix}$$

Apart from any interest such a proposal may have at descriptive and explanatory levels, one of its merits is that it provides a better fit with the analytic frameworks used in some other areas of language study. There is a parallel with some work in sociolinguistics. For example: Labov (1972) has analysed narratives of personal experience in black vernacular style as a clause sequence, and has also emphasized the role of conjunctions, simple subjects and adverbials. He views his narrative clause patterns as contrasting with ordinary conversation, however (1972:378), whereas the present data – white, middle-class vernacular – hardly justifies the need for such a contrast. As a further example, we may take child language acquisition, where most analyses of the young child credit him with cognitive or semantic discriminations involving location, time, frequency, and a whole range of deictic expression (*eg* Clark 1973). The importance of pronouns in relation to NPs is recognized (*eg* Limber 1976), as is connectivity, especially with *and* (*eg* Lust 1977, Crystal, Fletcher and Garman 1976:76). Adverbials are the main means of expressing spatio-temperal notions, and items such as *there*, *again*, etc are common in early samples. Later, adverbials are often used in the process of acquiring modal and other such structures (*eg maybe him go* for *he might go*). Given the recognition of these matters in this literature, which usually takes as its data-base domestic conversation, the contrast with standard adult grammars, which do not give them such emphasis, would be somewhat puzzling, without the hypothesis of the present paper. There is no evidence to support the view that the child in some way uses adverbials, pronouns or connectivity features less as he grows older. On the contrary, domestic performance, according to the present paper, stays very much the same, and a significant continuity can thereby be pointed out.

Looking at the data used to illustrate theoretical accounts of language, it has often been remarked that the examples cited are frequently somewhat contrived. This of course is inevitable if the aim of the exercise is to demonstrate the potential of language, for example defining the boundaries of grammaticality by repeated application of a set of rules until structures are generated which are wholly unacceptable. Lists of sentences of varying form, complexity and acceptability are the normal paradigms of illustration in linguistic writing. The value of this way of proceeding is undeniable, but it is a discovery procedure which, because of the way it is structured (involving an initial delimitation of a topic, and a systematic working

through of as exhaustive a range of formal permutations as one's ingenuity permits) is unlikely to encounter the data of spontaneous interaction. Because all such sentences – or at least most of them – are speakable, it is easy to imagine that there is no problem – that the grammar of informal domestic conversation is basically a reflection of that of the written language, with a few additional conventions such as ellipsis, intonation, and emphatic word order, and a few omissions, such as the structures characteristic of the more formal and literary modes of expression. The argument of the present paper, on the contrary, is that the linguistic organization of this variety of English has been fundamentally misconceived, due partly to the absence of data, partly to the uncritical application of traditional paradigms of enquiry. Considerable detailed descriptive work is now needed to take this claim further. Whatever the outcome, progress will be largely dependent on the use of the rigorous techniques of corpus-based analysis pioneered by Randolph Quirk. And with such precedent, an interesting outcome seems assured.