FACTORS INFLUENCING THE DISCOURSE FUNCTIONS OF EXISTENTIAL *THERE*
A CORPUS-BASED SURVEY

David B Umbach
Drake University

This paper is part of an ongoing large-scale study of the existential *there* construction, or ETC The full study, which involves over 4000 tokens gathered from a corpus of naturally occurring written texts totaling over 2 million words, is intended to address the following questions What is the actual function of the ETC in discourse? How is that discourse function affected by variations in the form of the ETC? "Discourse function" is seen here in relation to the ways that the ETC is used by writers to manage the flow of information in a text, "variations in form" are viewed in the context of grammatical complexity and embedding, both within the ETC and around it This paper primarily addresses questions that are subsidiary to the second question above First, what kinds of grammatical complexity are likely to be relevant to this study, and second, what kinds of grammatical complexity actually occur in ETCs in discourse?

For the purposes of this discussion, the ETC is defined as the structure that appears in the simple declarative as ***There* BE (as main verb) *NP***, as in *There is no Santa Claus* BE can occur in most, but not all, tenses and aspects (e g *there was, there has been,* but seldom if ever *there is being*) as well as in combination with any of the modals (*there can be, there will be, etc*) Also included are the corresponding question forms, *Is there a Santa Claus*, or *How many reindeer are there* (where the word *there* carries minimal stress when spoken), and non-finite forms such as *there having been a heavy rain* The noun phrase, usually referred to as the Post-Verbal Noun Phrase, or PVNP, to distinguish it from any of the canonical NP grammatical roles (subject, object, etc), can be very simple, as in *There is no Santa Claus*, a form sometimes referred to as a "bare" existential, or quite complex, as in Example 1, in which the PVNP is postmodified by a to-infinitive which contains a complex nominal *that* -clause, which itself comprises 3 clauses (matrix clause, a comparative and a conditional)

EXAMPLE 1.

There is abundant reason, nevertheless, to suppose that immaterial as these objections were, they would have been adhered to with a very dangerous inflexibility, in some States, had not a zeal for their opinions and supposed interests been stifled by the more powerful sentiment of selfpreservation (FP)
[Data Source Texts are listed separately after the References, alphabetized by abbreviation]

The above definition excludes tokens in which the verb is other than **BE**, as in *there arose a clatter* These are sometimes referred to in the literature as "presentational" rather than existential (see Birner and Ward 1998, pp 106-113 for a discussion of the functional contrast corresponding to the formal contrast in these two constructions) On the other hand, this definition does include tokens in which a transformational analysis would interpret the **BE** as the auxiliary in a progressive or passive verb phrase compare *there is a man in the corner* , corresponding to *a man is in the corner,* and *there is a man standing in the corner*, corresponding to *a man is standing in the corner* The existence of the participle form as post-modifier (*the man standing in the corner is the Dean*) suggests the categorization of these as tokens of *There* **BE** *NP* , and a preliminary survey of the data indicates that these tokens do in fact pattern functionally with the existential rather than the presentational

Previous work on the ETC has tended to concentrate on one of two issues the relationship of the ETC to semantically similar canonical sentences ("There is a man in the closet" versus "A man is in the closet") or the anomalous occurrence, in a small number of tokens, of definite NPs following the verb ("There's the Chinese restaurant down the street," in answer to the question "Where shall we eat?) Both types of study have shared, and to a certain degree confirmed, the assumption of traditional grammars that the ETC exists to introduce "new information" or "new participants" into a discourse (hence the anomaly of the definite NP, a form usually associated with "old" or "given" information), and neither type has given much (if any) attention to the potential for complexity within and around the ETC Both types of study have tended to rely heavily, if not exclusively, on constructed data, consisting almost exclusively of simple, monoclausal ETC's, and including, in those cases where the pragmatics of the ETC were at issue, equally simplistic specially constructed linguistic and situational contexts

Those studies that have dealt more effectively with the ETC, most notably Birner and Ward (1998), have more than adequately demonstrated that in order for an existential *there* sentence to be felicitous, the information represented by the post-verbal noun phrase (PVNP) must be "hearer-new" (Prince 1981), not just "discourse-new " Even these studies, however, have failed to sufficiently take into account the complexity of actual ETC's in discourse

Some forms of complexity are relevant here because of the nature of information "Information" is not a well-defined term, but whatever it is taken to mean, and whatever else may be said about semantic and pragmatic research centered on the concept of truth-conditions, it seems clear that one cannot discuss the status of an expression as information without acknowledging and accounting for those grammatical features that seem to be closely correlated to truth conditions (and related issues, such as presupposition) Such features include polarity and modality, and certain types of subordinate and embedded structures, such as conditionals and nominal *that* clauses

Other forms of complexity are relevant here because of the fact that previous discussions of the relationship between information status, sentence form and discourse structure have frequently based their definitions of discourse functional concepts like "cohesion," "topic," "focus" (and even "information") on patterns of inter-and intra-clausal connections between referring expressions (Halliday and Hasan 1976, for instance, or Martin 1992 381-492), without restrictions on the type of grammatical structure the referring expression occurs in The more complex a given token of a structure is, the more referring expressions it is likely to contain, and therefore, presumably, the more potential for such connections it will manifest Neglecting the complexity of the ETC means neglecting significant complications in what can or should count as "new," i e , "less connected" information

Thus we have at least two ways of discussing complexity that seem likely to be relevant to the goals of this project, the first involving factors that require more from a truth-conditional semantics than is required for a prototypical simple declarative sentence, the second involving factors that increase the number and variety of referring expressions (primarily noun phrases, but other constituents as well) within the construction Included in the first category are polarity, modality and the grammatical status of the ETC as a whole (i e , main clause, adverbial clause, relative clause) Included in the second are the number and type of structures found within the PVNP

These parameters of complexity have been, thus far, catalogued and analyzed for a subset of the corpus, consisting of 421 tokens (just over 10% of the total) As one example of the difference between corpus-based studies and constructed-data studies, consider the fact that while a substantial percentage of the previous work on the ETC deals with definite PVNPs (even Birner and Ward 1998 devote

approximately 75% of their discussion to definiteness in the PVNP), only 2 out of the 421 tokens in this sample, or 05%, have a definite PVNP In contrast, while little of the previous work has dealt with anything but main-clause ETCs (or existential *there* sentences), 129 tokens in this sample, or 30 6%, are subordinate• or embedded, not main-clause

The figures for polarity and modality are shown in Table 1

TABLE 1· POLARITY AND MODALITY

Negatives	135 (32%)
Modal Auxiliaries	31 (7%)

These figures include both clausal negation (*there was not*) and local negation of the PVNP (*there was no*), and all modal auxiliaries except those occurrences of *will* and *shall* that were clearly restricted to a "future time" reading Occurrences of *shall* with an "obligational" reading (*All of the people are equal under the law and there shall be no discrimination in political, economic or social relations because of race, creed, sex, social status or family origin* JC) are included

The most significant number here is the percentage of tokens that are negative, nearly a third of the ETCs in the sample Clearly negation must be accounted for in any complete discussion of the ETC in context Modality is somewhat less significant, although there are some examples where the "newness" of the information is largely contained in the modal Such cases also frequently involve a cohesive tie (Halliday and Hasan 1976 329-332) with a negative, as in Example 2

EXAMPLE 2.

The Dormouse again took a minute or two to think about it, and then said, `It was a treacle-well ' 'There's no such thing!' Alice was beginning very angrily, but the Hatter and the March Hare went `Sh! sh!' and the Dormouse sulkily remarked, `If you can't be civil, you'd better finish the story for yourself '
`No, please go on!' Alice said very humbly, `I won't interrupt again I dare say there may be ONE '   (AW)

The first ETC in this example is negative, and like many of the negatives in this sample is an explicit denial of "given" information, in this case, denial of the existence of a previously mentioned referent The new information in the second ETC is essentially just the replacement of the negation with the modality of possibility The referring expressions in the PVNPs (*such thing* and *one*) are anaphoric, and therefore by definition not new information

Similarly, in example 3, the new information in the second ETC replaces the negative polarity in the first with a possible modality (note that the negation in the first is not expressed within the ETC itself, but in the matrix clause within which the ETC functions as complement to the verb *think*)

EXAMPLE 3·

"Now, Mattie, I don't think for a minute that there are any Injuns in that Pass, and you must not be afraid We have got to go through it any way, but"--he hesitated,--"we may be mistaken, there

may be a few of them in there, and they'll have a mighty good chance to get in a shot or two "
(VA)

Again, the referring expression in the PVNP of the second ETC is anaphoric (*a few of them*),
and so by definition cannot be new information

Table 2 gives the numbers for a broad view of the grammatical status of the ETCs in the sample
Main-clause tokens do predominate, but as was noted above, more than 30% of the tokens here are not
main clauses Of the main-clause tokens, 282 (96 5%) are declarative, 11(3 75%) are interrogative

TABLE 2· GRAMMATICAL ROLE OF ETC

Main Clause	293 (69 5%)
Finite Adverbial Subordinate Clause	57 (13 5%)
Nominal *that* Clause	51 (12 11%)
Nominal Interrogative Clause	6 (1 4%)
Relative Clause	9 (2 0%)
Non-finite Verb Clause	2 (0 5%)
Part of Comparative Clause	3 (0 7%)

The most significant numbers here are the adverbial subordinate clauses (see Table 3, and
accompanying discussion, below), and the nominal *that* clauses Both are large enough classes to be
significant, and both forms can have special relationships to information status

The nominal *that* clause has a special status in regards to presupposition, as in Example 4

EXAMPLE 4·

And here it is all-important to notice, that, practically speaking, there can be no legal right to
resist the oppressions of the government, unless there be some legal tribunal, other than the
government, and wholly independent of, and above, the government, to judge between the
government and those who resist its oppressions, in other words, to judge what laws of the
government are to be obeyed, and what may be resisted and held for nought (ET)

The first of the ETCs in this example, in the first line, is a nominal *that* clause, complement to
the verb *notice* From a logical perspective, one cannot notice what does not exist, and therefore the
speaker (or in this case, writer) is presupposing the truth of the complement This particular example
also shows clearly one of the things that remains to be done in this study, this ETC is a negative
nominal *that* clause with a modal auxiliary Such overlapping complexities have not yet been
enumerated or analyzed

The second ETC in Example 4, starting on the second line, is an example of the majority of
adverbial subordinate clauses in this sample, i e , conditional clauses Table 3 displays the different
categories of adverbial subordinate clause types

TABLE 3 FINITE ADVERBIAL SUBORDINATE CLAUSE TYPES

Conditional	36 (63 0%)
Concessive	3 (5 0%)
Temporal	2 (3 5%)
Reason	7 (12 3%)
Result	1 (1 75%)
Cause	1 (1 75%)
Circumstance/Condition	5 (8 8%)
Alternative Conditional/Concessive	2 (3 5%)

In addition to being the most numerous, the conditionals are also the most interesting, for a variety of reasons For the general purposes of this study, of course, the subjunctive, hypothetical status of the conditional is significant to the question of information status However, in this sample, the conditional clauses have the added interest of being the focus of a possibly very significant variation in the distribution of ETCs based on genre

This sample includes three texts that could be labeled as "regulatory" or "constitutive," as opposed to more familiar genre labels like "narrative" or "argumentative " These texts are the Berne Universal Copyright Convention of 1988, the Constitution of Japan (1946-7) and the Treaty of the European Union In the sample as a whole, conditional clauses represent 63% of the adverbial subordinate clauses, which represent 13 5% of the total, that is, conditional clauses make up 8 5% of the total sample of 421 ETCs In these three texts, however, 10 out of 18 ETCs, or 55 6%, are conditional clauses Of the remaining 8, 3 are embedded within complex conditional clauses, and one is a main clause modified by a conditional Thus, 14 out of 18, or 77 8% of the ETCs in these texts are, in some sense, within the scope of a conditional

Furthermore, there are fewer ETCs in these texts than there are in the others (less than 1% of sentences in these three texts, versus 4 37% overall) These texts make far less use of the construction, and where it does occur, it is almost always within the scope of a conditional Additionally, many of the ETCs in these texts follow the pattern of example 5, in which most of the referring expressions (in this case, two out of three, *deficit* and *Member State*), while not anaphoric, have occurred numerous times in the preceding text

EXAMPLE 5·

The Commission may also prepare a report if, notwithstanding the fulfillment of the requirement under the criteria, it is of the opinion that there is a risk of an excessive deficit in a Member State (TU)

It seems reasonable to hypothesize, on this basis, that there is some sort of special relationship between the conditional and the ETC in this genre of texts As an informal characterization of this relationship, we can say that this genre of texts does not exist to convey information to the reader, but to direct the readers' actions based primarily on information presumed to be known The only information to be treated in such texts as hearer-new would be hypothetical, situations that do not, but may at some time, exist All information in the text, including that expressed in the ETCs, is actually hearer-old, the newness in the ETCs here is the conditionality imposed by the grammatical context

Moving on to the issue of internal complexity , table 4 shows the relationship of the ETCs in the sample to basic canonical sentence types (as categorized by Quirk et al 1985 1403-1408)

TABLE 4: INTERNAL STRUCTURE OF ETC

Bare Existentials	252 (60 0%)
SV	6 (1 4%)
SV passive	11 (2 6%)
SVC	16 (3 8%)
SVpassive C	4 (0 95%)
SVA	31 (7 4%)
NP + Relative Clause	57 (13 5%)
NP + Infinitive Clause	44 (10 5%)

Bare existentials predominate, but "bare" should not be confused with "simple " Example 5, above (repeated below as example 6a), for instance, is a bare existential, meaning that it cannot easily be paraphrased as a canonical sentence, without resorting to something like example 6b

EXAMPLE 6a

The Commission may also prepare a report if, notwithstanding the fulfillment of the requirement under the criteria, it is of the opinion that there is a risk of an excessive deficit in a Member State (TU)

EXAMPLE 6b

A risk of an excessive deficit in a Member State exists

This is not, however, a simple example Externally, it is a nominal *that* clause embedded within a conditional clause Internally, the PVNP contains three potential referring expressions the head noun and the NPs in each of the two post-modifying prepositional phrases The importance of such additional referring expressions will be discussed further below, in connection with Example 7

Table 5 presents a less complicated, and in some ways less problematic approach to the internal complexity of ETCs, a simple counting of the primary referring expressions, nouns and pronouns, in the PVNPs

This approach is less problematic in that, first, the categories are less ambiguous, and second, its implications are clearer, as far as previous studies of information management in discourse are concerned The categories of noun and pronoun are, in spite of difficulties one might face in constructing an accurate general definition, fairly clear-cut when it comes to tagging items in naturally-occurring discourse, especially written discourse Any item functioning nominally in a particular orthographically delineated sentence is counted as a noun, including onomatopoeic inventions such as "a long *sssssssss* " The implications of this approach are clearer simply because existing approaches to the study of information management, such as the "cohesion" and "texture" studies of Halliday and Hasan (19760 and Martin (1992), have traced patterns of information in texts primarily by examining patterns in the deployment of referring expressions, without noting any significant differences in the patterns dependent on the grammatical roles or relations thereof

TABLE 5: INTERNAL COMPLEXITY OF PVNP· Number of Noun and/or Pronouns within PVNP

Head Noun Only	81 (19 3%)
Head Noun + 1	124 (29 5%)
Head Noun + 2	70 (16 6%)
Head Noun + 3	54 (12 8%)
Head Noun + 4 or more	92 (21 8%)

As these numbers show, ETCs in which the head noun is alone in the PVNP do not predominate No single category listed here predominates The final category, Head Noun + 4, was chosen as the cutoff point simply because counting it in this way emphasizes the lack of predominance, above this point the numbers begin to drop off sharply, although there are, in fact, individual examples with as many as 17 nouns and pronouns, in addition to the head noun

The significance of even one additional referring expression in an ETC can be seen in example 7

EXAMPLE 7.

From the origin of San Francisco the earthquake has been its bane During the past fifty years fully 250 shocks have been recorded, while all California has been subject to them But frequency rather than violence of shocks has been the characteristic of the seismic history of the State, **there** having been few shocks that caused serious damage, and none since 1872 that led to loss of life

There was a violent shock in 1856, when the city was only a mining town of small frame buildings Several shanties were overthrown and a few persons killed by falling walls and chimneys **There** was a severe shock also in 1865, in which many buildings were shattered Next in violence was the shock of 1872, which cracked the walls of some of the public buildings and caused a panic (SF, emphasis added)

There are three ETCs in this passage (the word *there* has been put in boldface type, to assist the reader in locating them), and in all three, the head noun is the same word, *shock* In none of these can this word, in and of itself, be seen as new, given that appears twice in the 49 words preceding the first ETC

The second and third of these ETCs show more clearly the significance of the numbers in Table 5 Both of these are main-clause ETCs, so their information status is not affected by conditionality or external, grammatically imposed presuppositions The premodifiers in the two PVNPs are near synonyms, in this context, and so whatever newness there is to this information must be contained in the postmodification In both cases, the initial, restrictive postmodification resides in single prepositional phrases, which differ only in their complement NP The semantic contribution of these NPs to the PVNP as a whole is primarily responsible for the felicity of these ETCs

The primary goals of this paper have been to present data and discussion bearing on the question of variation in the form of ETCs in discourse, and the significance of that variation for an account of the discourse function of the ETC This much can be stated in conclusion ETCs occur in a wide range of

grammatical functions, as embedded constituents in complex sentences Many of these forms of embedding, such as conditional clauses and nominal *that* clauses, have long been recognized as bearing special relationships to meaning, especially within truth-conditional semantic frameworks, and therefore can be considered as, at least potentially, significant for the goals of the larger project in progress In addition, the form of the PVNP can be, and usually is, more complex than simply a single referring expression, and felicitous examples can be easily found in which the new information required for felicity is not represented in the head noun, but in other referring expressions located within modifying structures within the PVNP

REFERENCES

Birner, Betty J , and Gregory Ward 1998 Information status and noncanonical word order in English Amsterdam/Philadelphia John Benjamins

Halliday, M A K , and Ruqaiya Hasan 1976 Cohesion in English London Longman

Martin, J R 1992 English text system and structure Amsterdam/Philadelphia John Benjamins

Prince, Ellen F 1981 Toward a taxonomy of given-new information Radical pragmatics, ed by Peter Cole, 223-255 New York Academic Press

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik 1985 A comprehensive grammar of the English language New York Longman

DATA TEXTS CITED
(Abbreviation, Title and Author, followed by Project Gutenberg™ Filename, all data texts cited in this paper are available through Project Gutenberg, http //promo net/pg/)

AW Alice in Wonderland Lewis Carroll alice30 txt
CJ The Constitution of Japan (1946-7) jcnst10 txt
ET An Essay on the Trial By Jury Lysander Spooner tbjry10 txt
FP The Federalist Papers feder16 txt
SF The San Francisco Calamity By Earthquake And Fire Charles Morris, LL D sfclm10 txt
TU The Treaty of the European Union (the Maastricht Treaty) maast10 txt
VA Vanished Arizona Martha Summerhayes variz10 txt