

Coherence in Spoken and Written Discourse

Deborah Tannen, Editor
Georgetown University

VOLUME XII in the Series
ADVANCES IN DISCOURSE PROCESSES
Roy O. Freedle, *editor*



ABLEX Publishing Corporation
Norwood, New Jersey 07648

3 Coordination and Subordination Revisited: Syntactic Complexity in Spoken and Written Narrative Discourse*

Karen Beaman
Georgetown University

For speech is so much more than words and sentences.

John Steinbeck, *Travels with Charley*

It has generally been accepted that syntactic complexity in language is related to the number, type, and depth of embedding in a text. Syntactically simple authors use short, single clause sentences and rely more heavily on coordinated structures to provide cohesion and show relationships. Syntactically complex authors, on the other hand, use longer sentences and more subordinate clauses that reveal more complex structural relationships.

With regard to spoken and written differences in discourse, many linguists have found that written texts are syntactically more complex than spoken in that they exhibit a greater number of subordinate structures (O'Donnell 1974); spoken texts are syntactically simpler as they show a greater reliance on coordinate structures (Kroll 1977; Chafe 1982). Many authors have suggested reasons for this distribution. Chafe (1979) suggests this difference is due to the relationship between the speaker/writer and his or her audience. The speed of reading works with the slowness of writing to create an 'integrated' and 'detached' quality in written language as opposed to the 'involved' and 'fragmented' nature of speech.

*I would like to express my most sincere thanks to Deborah Tannen without whose help this paper would not have been written. In addition, I would like to thank Wallace Chafe, Tucker Childs, and Robin Lakoff for their valuable discussions and suggestions. However, any deficiencies are of course my own responsibility.

Lakoff (1979) suggests that the 'informality', 'spontaneity', and 'inconsequentiality' of speech in comparison to writing are some of the major factors influencing differences between spoken and written discourse. Working in the field of child language acquisition, Ochs (1979) introduces the concept of 'planning time' as an explanation for the greater use of coordinate or subordinate structures in discourse. With the extension of this concept to spoken and written differences in language, written discourse allows the writer greater time for planning, which results in increased complexity.

In opposition to what these investigators have found, Halliday (1979:47) asserts, 'contrary to what many people think, spoken language is on the whole more complex than written language in its grammar; and informal spontaneous conversation . . . is the most grammatically complex of all'. He suggests that this is because writing is static, whereas speech is dynamic. Poole and Field (1976) have found evidence supporting a similar hypothesis. They suggest that stylistics and the emphasis on simplicity and conciseness in writing inhibit the use of complex, elaborate structures in written discourse. These conflicting findings and hypotheses provide the motivation for the present investigation in the hope of uncovering what is really going on in spoken and written language.

This chapter reports the findings of a study that investigated differences in syntactic complexity between spoken and written narrative discourse in the attempt to better understand what processes are actually involved in each modality. The major findings show that the basic assumptions of many linguists in the past, i.e. that written language is more complex than spoken because it contains more subordinate and dependent structures, is too simplistic an assumption. The types and functions of sentence structures in discourse are so varied that an in-depth analysis, both formal and functional, is needed. The data show that if subordination implies complexity, then Halliday's hypothesis is in part right: spoken language is just as complex as written, if not more so on some measures. The differences between spoken and written language that many linguists in the past have found may be more a result of differences in the formality and purpose of the discourse than in a spoken/written distinction. Evidence for these hypotheses will be given in the detailed analysis that follows of the types and functions of sentence structures found in spoken and written narrative discourse. This study suggests a methodology for studying syntactic complexity in discourse and provides a broader basis for understanding differences between spoken and written discourse in general.

BACKGROUND

Kroll (1977), in an article on subordination and coordination in spoken and written narratives, provides evidence confirming the hypothesis that written language uses more subordinate structures whereas spoken uses more coordinate

structures. Her data show that written narratives contain 35 percent subordinate/dependent structures, whereas spoken narratives contain only 14 percent. For coordinate structures, written narrative shows only 25 percent, but spoken shows 40 percent.

This high percentage of coordinate structures in the spoken narratives is attributable largely to a high percentage of occurrence of the coordinator *and*. However, as Kroll says herself, in spoken discourse *and* (often pronounced [ən] or [n]) functions other than simply as a coordinator; primarily, it serves as a filler word, to hold the floor for the speaker to indicate that he or she is about to say something more (Kroll 1977:95). When Kroll considers coordinate structures without *and*, only 4 percent of the clauses in the spoken are coordinated: 96 percent of her coordinators are *and*. Such a high percentage of a single feature in a text leads the wary investigator to suspect that there may be other processes at work influencing the distribution. It may not simply be the case that spoken discourse contains more coordinate clauses and written more subordinate clauses. In order to uncover what these processes actually are, it is important to look at the types and functions of subordinate and coordinate clauses in the two modes of discourse.

Kroll presents an analysis of the types of subordinators, coordinators, and phrases found in her data, but her tables present only absolute numbers rather than percentages or frequencies of occurrence. Consequently, the data are largely uninterpretable because of the varying lengths of each narrative. Because absolute numbers naturally increase with the length of each text, a frequency index would be a more representative measure.

In a study aimed at showing the greater complexity of writing over speech (as determined by sentence length and depth of embedding), O'Donnell (1974) has found that the average length of syntactic units in written language is greater than in spoken language. Using the T-unit identified by Hunt (i.e. 'one independent clause and the dependent clauses [if any] syntactically related to it' [O'Donnell 1974:103]), O'Donnell assumes that this greater average length is indicative of greater syntactic complexity in writing, 'a condition associated with compactness in verbal organization' (p. 108). O'Donnell also found a significantly greater number of dependent clauses in written discourse. Gerunds, participles, attributive adjectives, passive constructions, and modal and perfective auxiliaries were more frequent in his written sample, whereas nominal clauses, infinitives, and progressive auxiliaries were more common in the spoken.

Chafe (1979, 1982), working in the field of cognitive linguistics, proposes an explanation for why syntactically complex structures occur more frequently in written discourse. He suggests that this greater dependence on subordinate structures in writing is due to the relationship of the writer to his or her audience. As writer and audience are generally removed from each other in time and space, writing has a 'detached' quality as opposed to an 'involved' quality found in speech. In addition, the slowness of writing and the speed of reading allow the

TABLE 1
Chafe's (1979, 1982) Features of Spoken and Written Language

Integration	Involvement
Nominalizations	First person references
Past/present participles	Speakers' mental processes (e.g. 'I think')
Attributive adjectives	Monitoring information flow
Conjoined phrases	Emphatic particles (e.g. 'just', 'really')
Series (of noun/verb phrases)	Fuzziness, vagueness, hedges
Sequences of prepositional phrases	Direct quotes
Complement clauses	
Relative clauses	
Introductory phrases (e.g. 'In view of this')	
Adverbial conjunctions	

Detachment	Fragmentation
Passives	Coordinating conjunctions
Nonactive subjects	Afterthoughts
Precision	Repetitions
Indirect quotations	False starts

writer time to 'integrate' his or her ideas into a more complex, coherent whole. Opposed to this is the 'fragmented', spurtlike nature of speech in which ideas are loosely joined by *and*.

Specific devices that reveal each of these features are detailed in Table 1. These four features—integration, involvement, fragmentation, and detachment—form the poles of two intersecting continua on which various types of discourse can be plotted. Putting integration-fragmentation on the vertical axis and detachment-involvement on the horizontal axis, Chafe (1979) suggests that the prototypical sample of speech falls in the lower right quadrant—highly fragmented and involved—whereas the prototypical sample of writing falls in the upper left quadrant—highly integrated and detached.

In light of this research, Tannen (1980, 1982) suggests that written fiction combines the involvement factor of speech and the integration factor of writing, placing the prototypical example in the upper right quadrant of Chafe's grid. Tannen explains this by reference to oral and literate strategies. Written fiction, as opposed to other types of written language, uses strategies that have been associated with oral tradition: that is, 'it builds upon the immediacy function of spoken language—"imageability" and "involvement"', (Tannen 1980:214).

In the same vein as Chafe's work, Robin Lakoff (1979) argues that the accepted dichotomy between spoken and written language is a false dichotomy; there is, rather, 'a continuum of discourse, arranged as to the purpose of the discourse and the environment in which it occurs' (Lakoff 1979:23). She identi-

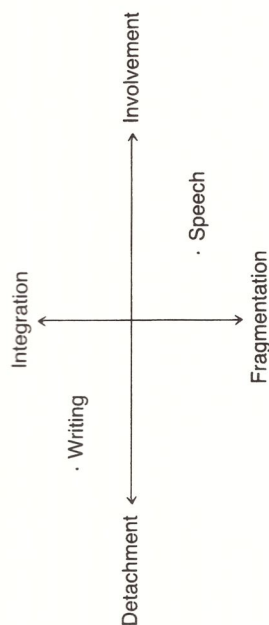


FIGURE 1. Schematic Representation of Features of Spoken and Written Language (from Chafe 1979).

fies six criteria as a set of binary features to be used in distinguishing different types of discourse along the continuum: 'visibility', 'reciprocity', 'informality', 'spontaneity', 'empathy', and 'inconsequentiality'. At the poles of this continuum are the spontaneous oral dyad and expository written prose. The former exhibits all pluses for these six features, whereas the latter shows all minuses (Table 2). Other types of discourse, such as telephone conversations, lectures, letters, and psychotherapeutic discourse, fall on the continuum between these poles.

Though not focusing directly on differences in spoken and written discourse, Ochs (1979) suggests the idea of evaluating the amount of 'planning time' that goes into a text. 'Planned discourse is discourse that has been thought out and organized (designed) prior to its expression'. 'Unplanned discourse is discourse that lacks forethought and organizational preparation' (Ochs 1979:55). Ochs describes four features that she says are characteristic of unplanned discourse: (a) reliance on immediate context to express propositions; (b) reliance on morphosyntactic structures acquired early in life; (c) tendency to repeat and

TABLE 2
Lakoff's (1979) Features of Spoken and Written Language Making up the Discourse Continuum

Feature	Spoken (oral dyad)	Written (expository prose)
	+	-
Visibility	+	-
Reciprocity	+	-
Informality	+	-
Spontaneity	+	-
Empathy	+	-
Inconsequentiality	+	-

replace lexical items in the expression of a proposition; and (d) similarity in the form and content of sequentially arranged social acts.

The second feature is important for an analysis of syntactic complexity in discourse. Unplanned discourse exhibits syntactic structures similar to those used by children in the early stages of language development. Developmentally, younger children more than older children have been found to use more coordinate structures (rather than subordinate), determiner plus noun constructions (rather than relative clauses), deictic modifiers (rather than definite articles), active voice (rather than passive), and present tense (rather than past or future tenses).

Ochs suggests that linguistic development gradually progresses from pragmatic devices for connecting utterances in discourse to greater reliance on syntax. Likewise, the more planning put into discourse, the greater the move from pragmatic devices and reliance on immediate context to more complex syntactic structures for providing cohesion and showing relationships between ideas.

Because written discourse allows the writer more time to structure his or her ideas, it will naturally be more planned than its spoken counterpart. Thus, it follows that written discourse should be syntactically more complex than spoken discourse and exhibit a greater reliance on subordinate clauses.

Directly opposed to the findings of these studies, however, Halliday (1979) suggests that speech rather than writing is the more complex mode of discourse. His hypothesis is based on the degree of 'lexical density' in a text. He suggests that speech shows a low lexical density in that it uses fewer high content words but more clauses, whereas writing exhibits a high lexical density (cf. O'Donnell's 'compactness'), showing more high content words but fewer clauses. As he puts it, 'speech has complex sentences with simple words, while writing has complex words in simple sentences'. Thus, the types of complexities involved in each mode are of very different types (i.e. complex sentences versus complex words). These are not discrete phenomena but rather tendencies in certain directions. It is the degree of what Lakoff calls 'spontaneity' of the text that influences the type and degree of complexity.

Evidence supporting Halliday's hypothesis can be seen in the work of Poole and Field (1976). With regard to syntactic complexity, they found that though average sentence length is longer, verbal complexity greater, and the use of adjectives more frequent in their written samples, spoken discourse shows a significantly greater degree of subordination, elaboration of syntactic structures, and use of adverbs (at the .01 significance level or better).

Poole and Field suggest that the greater syntactic simplicity found in written discourse is a direct product of stylistics and the emphasis on simplicity, conciseness, precision, and clarity in writing classes. In speaking, they find no such 'economy conciseness corrective feedback' inhibiting the discourse (p. 309). Because of this 'inconsequential' (cf. Lakoff) nature of speech, speakers do not feel the need to monitor themselves to the degree that is required in

writing, where words are of a more permanent nature. This seems to be an important factor influencing the complexity of discourse.

One reason Kroll, O'Donnell, Chafe, Poole and Field, and others may have obtained such seemingly significant, yet divergent, results is that their samples of spoken and written language are not entirely comparable. Their samples lay at opposite ends of Lakoff's formality/spontaneity, or Ochs's planned/unplanned continuum. All four studies have hopelessly confounded two crucial features of discourse: modality (i.e. spokenness/writtenness) and register (i.e. purpose or formality). Because of this problem, what looks like differences between spoken and written discourse may really be differences in the register, purpose, formality, or amount of planning time of each task.

Kroll (1977) analyzes spontaneous spoken narratives, in which subjects recounted an episode in which they were in danger of death (relatively unplanned/informal) with the same narratives written overnight as a composition assignment (highly planned/formal).¹ O'Donnell (1974) compares spontaneous spoken language from a television talk show (relatively unplanned/(in)formal?) with published newspaper columns (highly planned/formal). Chafe's samples are taken from informal dinner conversations (highly unplanned/informal) and published academic papers (highly planned/formal).² And Poole and Field (1976) contrast spoken samples taken from interviews with students concerning their university experiences (relatively unplanned/informal) with written narratives obtained in a 'group testing situation' on the students' 'life-forecasts' (highly planned/(in)formal?).

It seems obvious that the samples of language used in these studies not only contrast in being spoken and written, but also in their register and purpose: the spoken are all relatively unplanned, informal, and inconsequential, whereas the written are generally planned, formal, and full of consequence. Clearly, these differences in register, purpose, and degree of formality/planning must be controlled if the real influences on spoken and written language are ever to be determined.

The present study uses a corpus of spoken and written narrative discourse that has been controlled for such differences in register, purpose, and formality. Tannen (1982) suggests that narrative discourse is different enough from other types of discourse to constitute its own distinct genre. Indeed, Labov's work on discourse has shown narrative to be markedly different from other types of conversation or writing. Labov (1972:377) observes that 'narrative as a whole contrasts sharply with ordinary conversation, which shows a much more complex structure'. This storytelling mode, which speakers and writers naturally

¹It might be mentioned here that in a controlled experiment Golub (1969) found that the order in which samples are taken from subjects (i.e. whether they write first and then speak or speak first and then write) has no significant effect on the results of the experiment.

²See Chafe (1982) for an analysis that looks into this formality distinction.

TABLE 3
Lakoff's Criteria Applied to the Present Corpus of Spoken and
Written Narrative Discourse

Feature	Spoken Narrative	Written Narrative
Visibility	+	-
Reciprocity	-	-
Informality	+	+
Spontaneity	+	+
Empathy	+	+
Inconsequentiality	+	+

adopt when telling a narrative, may be largely exempt from the traditional models and structures of ordinary spoken conversation and of expository written prose, requiring a need for the development of very different models and structures for analysis. When applied to narrative discourse (with particular reference to the corpus of the present study), Lakoff's criteria for determining various discourse types along the continuum reveal the characteristics shown in Table 3.

Clearly, oral and written narratives do not differ in the way that typical samples of spoken and written discourse do (i.e. the oral dyad versus expository prose, cf. Table 2). Narratives are by definition nonreciprocal because the narrator holds the floor for the duration of the story. Though this is not necessarily true of all narratives, the spoken and written narratives in the present study are identical in being relatively informal, spontaneous, generally empathic, and of a relatively inconsequential nature.

The only real difference between the two discourse types is 'visibility', the most basic criterion distinguishing spoken and written language. Lakoff (1979:7) points out that 'visibility gives speakers extra channels through which to communicate'. Extralinguistic signals (e.g. eye contact, gestures, stance, distance) and many of Chafe's features of involvement and fragmentation (Table 1) (e.g. first person references, evidence of speakers' mental processes, monitoring information flow, false starts), which carry considerable additional information for the listener, are simply not available to the writer due to the lack of visibility.

Based on Lakoff's criteria, the narratives in the present study provide an ideal corpus for the investigation of syntactic complexity in spoken and written discourse because the two discourse types are comparable in all major respects except the one under investigation: one is spoken, whereas the other is written.

THE STUDY

The data for this analysis consist of 20 spoken narratives and 20 written narratives about the 'pearl film', a short movie commissioned for a project in 1975 by

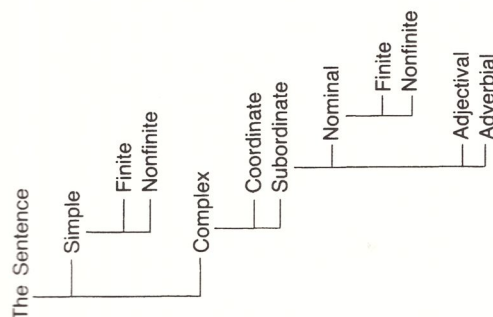


FIGURE 2. Schematic View of Sentence Structure (adapted from Quirk et al. 1972).

Wallace Chafe and his associates at the University of California, Berkeley.³ The purpose of the film was to elicit natural, spontaneous discourse from multiple speakers on the same topic. After viewing the film, subjects were asked to tell what they saw, either in speech or writing. All 40 narratives were told by university women, under similar circumstances. (Speakers told their narratives one at a time; writers wrote simultaneously.) Because subjects were not told ahead of time the task they were to perform, the data represent examples of unplanned, informal, spontaneous narrative discourse.

Due to the varying lengths of narratives, a frequency index is used to measure the occurrences of particular features. A simple frequency index is calculated by dividing the total occurrences of a particular word/feature/structure by the total words in the narratives (12,594 for the 20 spoken stories; 7072 for the 20 written stories) then multiplying the result by 1000. This yields an index that is interpreted as the number of occurrences of that structure per every 1000 words.

Schematically, the view of the sentence adopted here is a combined formal and functional approach as depicted in Figure 2, based partly on Quirk et al. (1972), *A Grammar of Contemporary English*. (Much of their detail has been left out in the interest of simplicity and generalization.)

Basically, sentences can be viewed as being of two types: simple or com-

³These narratives also provide data for Tannen and Michaels & Collins (this volume) as well as Bernardo (1979). Transcripts of the spoken narratives appear in Chafe (1980). Copies of the written narratives may be obtained from Tannen. See preceding chapter, p. 23, for a summary of events in the film.

plex. Simple sentences obligatorily consist of only one finite clause and may have any number of nonfinite clauses attached. Sentences that are complex are either coordinate or subordinate. Functionally, subordinate clauses can be either nominal, adjectival, or adverbial, with nominal clauses being, structurally, either finite or nonfinite. The following analysis will proceed along these lines in defining and describing the types of sentence structures found in spoken and written narrative style discourse.

THE SENTENCE

Sentences in English are of two basic types: simple (i.e. containing one clause) or complex (i.e. containing more than one clause). With a surface structure definition, a clause can be seen as a structure with a finite verb (i.e. showing tense, mood, aspect, and voice) and its constituent parts (i.e. subject, objects, adverbs, etc.). Nonfinite structures are those not showing tense or mood but still indicating aspect and voice (Quirk et al. 1972:43).

Though not considered in the analysis of sentence structure by Quirk et al., fragments are a structure found in spoken and informal discourse. In this study, fragments are defined as finite subjectless sentences standing alone (i.e. without a main clause) in the discourse, for example:⁴

EW17:35 Bruises chin and spills pears.

EW6:13 Back to pear-picking shots & tree cuts for awhile.

Simple sentences contain one and only one finite clause and optionally may contain many additional nonfinite phrases. In the following examples, the finite verb in the sentence is underscored.

Simple Finite Sentences (SFS)

ES2:93 And then he crashes into a rock.

ES6:66 Then they come across his hat on the road.

EW1:6 Then a boy rode up on a bicycle.

EW5:37 The man leaned back against the tree in puzzlement.

Simple Sentences with Non-Finite Phrase(s) (SNFS)

ES6:4 It opens with I guess a farm worker picking pears in a tree.

ES12:132 Upsetting his bicycle, upsetting the pears, she rides past.

⁴In these and all following examples: ES stands for English Spoken narrative, EW for English Written narrative; the number before the colon stands for the narrative number (1–20), the number after the colon for the clause number within the narrative.

EW2:1 The film begins with a rather portly Mexican picking green pears.

EW12:38 The boy pulled up his pants leg in order to rub his shin.

Whereas simple sentences consist of only one finite clause, complex sentences, by definition, contain more than one finite clause. There are two major devices for linking clauses of complex sentences: coordination and subordination. Coordination is the symmetrical relation in which two or more elements of the same status are linked together. Subordination is the nonsymmetrical relation holding between two clauses such that one is a constituent or part of (i.e. dependent upon) the other (Quirk et al. 1972:720); there is a hierarchy in the relationships. Thus coordination links two or more independent clauses, whereas subordination joins one independent clause with any number of dependent clauses.

Coordinate Sentences

EW7:43 The boy with the bike started out again & the 3 boys walked on in the opposite direction.

ES2:17 And a man comes along with a goat, and the goat is obviously interested in the pears.

Subordinate Sentences

EW7:45 The boy who had been playing paddle ball noticed the straw hat on the ground.

ES10:39 As he's doing this, somebody walks by with a goat or something.

Each coordinate sentence above consists of two independent clauses, each capable of standing as a separate sentence on its own. The subordinate sentences consist of one independent clause and a dependent clause that is subordinated to it. These subordinate clauses cannot stand as sentences on their own (e.g. 'who had been playing paddle ball').

Coordinated clauses may occur in a series, and thus be marked with only one overt coordinator. They also allow ellipsis of the subject of the clause they introduce if the subject is coreferential with that of the preceding clause (Quirk et al. 1972:555).

EW12:10 The man wiped some pears with the scarf, (Ø) tied it back on, and (Ø) climbed the ladder again.

ES3:40 So one of the boys whistles to him, and (Ø) stops him, and (Ø) gives him his hat back.

Each of these two sentences is made up of three coordinated clauses, though in the first sentence only the last clause is explicitly introduced by *and*, and the second two clauses of each are subjectless—indicated by (Ø).

A subordinate clause has a finite verb and is usually introduced by a subordinating conjunction. (The only exceptions to this definition are non-finite clauses, the *zero that*-clause, and *contact relatives* which will be treated later.) Subordinate clauses cannot stand without a main clause of which they form a constituent part.

EW12:27 A few pears fell from the basket as he hit a couple of bumps.

ES6:16 And you see a guy leading a goat past the tree where he's picking the pears.

The major semantic difference between coordination and subordination is that the information is asserted in the former but presupposed in the latter (Quirk et al. 1972:551). Consider the following hypothetical examples.

- (1) The baskets were full, but the pear-picker kept picking pears.
- (2) Although the baskets were full, the pear-picker kept picking pears.
- (1) His pockets were full and he emptied the pears into the basket.
- (2) Because his pockets were full, he emptied the pears into the basket.

The cause-result relationship is the same for the two sets of sentences, but in (1), 'his pockets/baskets were full' is stated (i.e. asserted), while in number (2) it is assumed (i.e. presupposed).

THE ANALYSIS OF SYNTACTIC STRUCTURE

Table 4 presents an overview of the types of sentences that occur in the spoken and written pear narratives based on the preceding taxonomy of sentence types. In this and all subsequent tables, the numbers in parentheses represent the absolute number of occurrences of that structure in all 20 narratives.

As would be expected, fragments occur more often in speech than in writing. Because of the warnings of writing teachers against using fragments, it is surprising that even 3 percent of the sentences in the written stories are fragments. Of the ten that occurred, five were used by one speaker and three by another. Thirty-seven percent of the first writer's sentences were simple sentences without nonfinite clauses, and only one of her sentences had a subordinate clause. These fragments, combined with the heavy reliance on simple sentences, stylistically produce a staccato effect. Tannen (this volume) and others have called this 'telegraphic style'.

EW17:20 Decides not to—leaves—returns.
Picks up whole basket while the picker is up in tree.
Puts on the front of his bike and rides away.

TABLE 4
Types of Sentence Structures in the Spoken and Written Pear Narratives

	Spoken (see footnote 5)		Written
Fragments			
Simple:	Nonfinite (SNFS)	9% (57)	3% (10)
	Finite (SFS)	6% (37)	11% (46)
Complex:	Coordinate	20% (136)	18% (74)
	Subordinate	25% (168)	38% (158)
	Coor/Sub ^a	13% (88)	12% (50)
		27% (178)	18% (76)
Total		100% (664)	100% (414)

^aCoor/Sub^a stands for coordinated and subordinated clauses.

Table 4 shows that simple sentences with accompanying nonfinite clauses (SNFS) occur less often in both modes of discourse than simple finite sentences (SFS). These SNFS are more common in the written narratives (but only 5 percent more) than in the spoken. This fact may reflect the greater syntactic complexity of writing as the addition of nonfinite clauses to the simple finite sentence seemingly requires increased mental processing by creating greater lexical density (cf. Halliday).

The complex coordinate sentence is the most common type of sentence structure in the narratives—25 percent in the spoken and 38 percent in the written. Contrary to what Chafe (1982) and Kroll (1977) found, there are significantly more coordinated sentences in the written stories (13 percent more) than in the spoken. Thus, when looking at sentence types and not simply at the frequency of occurrence of a particular word (i.e. *and*, as Chafe and Kroll did), those instances of *and* that serve merely as filler words can be eliminated, and the overall picture of true coordination in discourse is revealed.

Also contrary to the findings of Chafe and Kroll, the percentage of subordinated sentences is low in both the spoken (13 percent) and written (12 percent) narratives. The percentage of subordination in discourse used as a measure of greater syntactic complexity, clearly shows that neither the spoken nor written narratives can be considered the more complex mode of discourse.

The Coor/Sub construction, that is, some combination of both coordinate and subordinate clauses in one sentence and seemingly a more complex construction, is more frequent in the spoken stories (8 percent more). The following sentences are examples of the Coor/Sub construction.

EW1:19 As they passed he turned to look at her, and his hat blew off.
ES3:7 And the guy who is picking pears, picks the pears and puts them in these baskets that he has.

TABLE 5
Number of Finite Clauses Joined by Coordination in the Spoken and Written Pear Narratives

<i>Number of Coordinate Clauses</i>	<i>Spoken</i>	<i>Written</i>
2	44% (152)	65% (152)
3	22% (76)	22% (52)
4	15% (51)	9% (22)
5	5% (16)	2% (4)
6	4% (15)	1% (2)
7	6% (21)	1% (2)
8	2% (6)	— (0)
9	.3% (1)	— (0)
10	.3% (1)	— (0)
11	1% (5)	— (0)
12	.3% (1)	— (0)
13	.3% (1)	— (0)
Total	100% (346)	100% (234)

That this complex construction is found more frequently in the spoken stories seems to support Halliday's argument for the greater complexity of speech. A look at the number of clauses within these sentences will perhaps provide a better understanding of these differences.

The number of finite clauses within each sentence is presented in the following two tables. These tables present a breakdown of the last three sentence types (i.e. only the complex sentences) from Table 4. Table 5 shows the number of finite clauses from the purely coordinate group and from the coordinate/subordinate group joined by some type of coordination.

As can be seen from Table 5, the written narratives show 21 percent more sentences with only two coordinated clauses. Coordinating three clauses together in one sentence occurs with the same frequency in both the spoken and written stories. Perhaps most noteworthy here is the greater tendency in the spoken narratives to join numerous clauses into one sentence. One speaker conjoins thirteen clauses into a single sentence!⁵ This chaining effect reflects the sequential nature of narrative style discourse. The following example typifies this feature of spoken discourse.

⁵The criteria used to distinguish the boundaries of a sentence in spoken discourse are based on Chafe's (1980:26) concept of an 'intonational sentence', that is, a sequence of idea units typically bounded by sentence-final intonation and syntactic closure, expressing a single 'center of interest'. These decisions were made by workers on the Chafe project, and I have accepted their transcriptions.

ES1:20 And then he gets down out of the tree,
and he dumps all his pears into the basket,
and the basket's full,
and one of the pears drops to the floor,
and he picks it up,
and he takes his kerchief off,
and he wipes it off,
and places it in the basket
which is very full.

It is clear that the *ands* in this example, particularly those initiating clauses, are serving as filler words or weak connectives paralleling the use of the adverb *then* (see Table 18 and discussion). This is a common feature of narrative discourse, as the very nature of narrative consists of the sequential retelling of events.

Table 6 shows the number of finite subordinate clauses making up the complex sentences from Table 4. Again, both the spoken and written stories show a preference for embedding only one subordinate clause per sentence. The written narratives show a slightly greater tendency to embed only a single subordinate clause (8 percent more), whereas the spoken show a slightly greater tendency to embed multiple clauses within a single sentence, some speakers using four or five subordinate clauses per sentence. But overall, numerous subordinate clauses are not as common to either modality as are numerous coordinate clauses (Table 5).

The results in Table 6 are contrary to O'Donnell's findings. O'Donnell found more single-clause units in speech, and more multi-clause units in writing. With the assumption that the greater number of embedded clauses within a sentence is evidence of increased complexity, the spoken narratives appear more complex than the written. This would confirm Halliday's hypothesis that spoken

TABLE 6
Number of Finite Clauses Joined by Subordination in the Spoken and Written Pear Narratives

<i>Number of Subordinate Clauses</i>	<i>Spoken</i>	<i>Written</i>
1	71% (187)	79% (99)
2	20% (53)	18% (23)
3	6% (17)	3% (4)
4	2% (5)	— (0)
5	1% (1)	— (0)
Total	100% (263)	100% (126)

language uses more complex clauses than written with reference to narrative discourse.

In summary, then, Tables 4, 5, and 6 show that although there is a greater overall percentage of coordinated sentences in the written narratives (Table 4), the spoken stories show a greater tendency to string numerous clauses together into a single sentence (Tables 5 and 6). It has also been suggested that merely counting instances of the word *and* presents an inaccurate picture of coordination in discourse, because the primary function of *and* (i.e. as a weak connective and filler word) has no counterpart in the written modality. The linguist should look at the function of the sentence as a whole as a more revealing and reliable means for determining syntactic structure. In order to further demonstrate this point, the following section will present a frequency analysis of the types and functions of coordination and subordination found in the spoken and written pear narratives.

COORDINATION

And, *but*, and *or* are the three main coordinating conjunctions. Some examples follow.

EW10:39 The boys didn't say a word, but simply walked past munching their pears.

ES9:41 And I don't know how he does it, but he puts the big basket of pears on his bicycle.

So, *for*, *yet*, and *nor* can also occur as coordinators, though there are no instances of the latter three in the data. Table 7 shows the frequency indices for coordinators in the spoken and written pear stories.

Contrary to the foregoing analysis of actual sentence types, which showed coordinate sentences to appear with a greater percentage in the written narratives (cf. Table 4), the actual frequency of occurrence of coordinating conjunctions is over twice as great in speech (Table 7). This apparent anomaly can be resolved by reference to the different methods and purposes of each analysis. Whereas the previous analysis provides an overall picture of the use of coordination as a device for combining and structuring sentences, Table 7 presents an absolute frequency of occurrence resulting from a 'blind' counting of each instance of a particular word.

The method of absolute frequency has been used in many previous studies and has promoted the assumption of the greater frequency of coordinating conjunctions in speech than in writing. Based on this method of analysis, the heavy use of coordinating conjunctions in the spoken narratives has been described by Chafe (1979, 1982) as the 'fragmented' nature of speech.

However, as has been pointed out, not all instances of *and* in the spoken data constitute true coordination. The instances of *and* in the spoken examples

TABLE 7
Frequency Indices for Coordinating Conjunctions in the Spoken
and Written Pear Narratives

	Spoken	Written
and	72.9 (918)	35.9 (254)
but	4.8 (61)	2.1 (15)
so	4.2 (55)	.8 (6)
or	1.1 (14)	.1 (1)
and so	1.3 (16)	— (0)
Total	84.5 (1064)	39.0 (276)

above show that it more often serves as a filler word, a weak connective, as the speaker is planning his or her next utterance. This use of *and* contributes to the fragmented quality of speech. This greater use of filler words and the characteristic of chaining numerous clauses together with *and* can be attributed to speakers' lack of tolerance for silence. The tendency in speech is to fill up empty space and not leave time for an awkward silence to build up.

When a single word or feature is used with such an enormous frequency as these coordinators (84.5), it diminishes in its semantic and functional load, taking on a more generalized meaning. Martinet (1964:177) points out, 'there is a constant and inverse relationship between the frequency of a unit and the information which it conveys'. It appears that this is the case with *and* in spoken narrative. Its use as a filler word, as an introducer of clauses, and as an indicator of the sequential ordering of events in a narrative (see Table 18 and discussion), has so generalized its meaning that it has lost its specific coordinating function.

SUBORDINATION

As has been stated, subordination is the asymmetrical relationship between an independent and dependent clause(s) in which the dependent clause is introduced by an overt subordinating conjunction. Subordinate clauses can be divided into three types with respect to their function in the whole sentence: nominal, adjectival, and adverbial. The following discussion analyzes the subordinate clauses in the pear stories in terms of these sentence types.

Nominals

Nominal subordinate clauses are those clauses that function as noun phrases. Nominal clauses fall into five major categories: (a) *that*-clauses, (b) interrogative clauses, (c) nominal relative clauses, (d) *to*-infinitive clauses, and (e) *-ing*

clauses (Quirk et al. 1972:734). The first three are finite nominal clauses (i.e. showing tense, mood, aspect, and voice); the latter two are nonfinite nominals (i.e. showing only aspect and voice). This proves to be an important distinction.

That- Clause

That-clauses, or complementizers, are the most common type of nominal subordinate clause. When the *that*-clause is not functioning as the subject of the clause it introduces, the conjunction is frequently omitted in informal usage, leaving the zero *that*-clause [indicated by (Ø)] (Quirk et al. 1972:734).

ES2:142 He doesn't even notice that the pears are stolen yet.

EW5:9 I think that someone rode or walked by at this point.

ES2:110 I think (Ø) his ego was hurt.

EW4:35 One of the three notices (Ø) the boy's hat is on the dirt still.

Table 8 presents the frequency indices for nominal subordinate *that*-clauses in the spoken and written pear narratives. As the findings show, *that*-clauses occur with almost twice the frequency in the spoken narratives as in the written. This is contrary to what Chafe (1982) found in written discourse as an indicator of the 'integrated' nature of writing. His findings show frequency indices of 5.1 for spoken discourse and 8.0 for written. However, as has been pointed out, Chafe's samples were of informal dinner conversations and expository prose: opposing ends of the formality continuum.

Wh- Interrogative Clause

Wh-interrogative clauses, the second type of nominal subordinate clause, resemble *wh*-questions in that they leave a gap of unknown information.

ES2:55 You wonder how he's going to take it

TABLE 8
Frequency Indices for *that*-Complementizers in
the Spoken and Written Pear Narratives

	Spoken	Written
that	4.5 (56)	2.7 (19)
(Ø)	2.1 (27)	1.1 (8)
Total	6.6 (83)	3.8 (27)

TABLE 9
Frequency Indices for Nominal *wh*-Interrogative
Subordinate Clauses in the Spoken and Written
Pear Narratives

	Spoken	Written
how	1.0 (12)	— (0)
what	1.0 (12)	— (0)
why	.1 (2)	— (0)
Total	2.1 (26)	— (0)

ES9:73 I don't know what you call them

ES2:123 I thought, why didn't he think of it before

Table 9 shows the frequency indices for nominal *wh*-interrogative subordinate clauses in the spoken and written pear stories.

It is interesting to see that there are no instances of *wh*-interrogatives in the written stories. Because these nominal clauses resemble questions, one would not expect to find them in written discourse (except of a rhetorical nature) due to the lack of the 'visibility' characteristic of written language. Their occurrence in the spoken stories may be a result of the more personal nature of spoken language (cf. Chafe's evidence of speaker's mental processes in spoken discourse). This feature of spoken discourse lacks a comparable feature in informal written usage.

Nominal Relative Clause

Nominal relative clauses, the third type of nominal subordinate clause, function more like noun phrases than other nominal subordinate clauses. They can usually be paraphrased by a noun phrase containing a postmodifying relative clause (Quirk et al. 1972:738).

ES5:10 I guess what he's picking is pears
(i.e., I guess it's pears that he's picking)

ES2:131 You just know that those little boys are going back to where the pear
tree was.
(i.e., ... going back to the place that the pear tree was)

Table 10 shows the frequency indices for these nominal relative clauses in the spoken and written stories. Again, there are no examples of nominal relative clauses in the written narratives. The examples seem to suggest that speakers use this structure to monitor the needs of the listener and provide the right amount of information (cf. Chafe's involvement factor and Lakoff's empathy criterion).

TABLE 10
Frequency Indices for Nominal Relative
Subordinate Clauses in the Spoken and Written
Pear Narratives

	Spoken	Written
what	1.3 (16)	— (0)
where	.6 (8)	— (0)
whatever	.2 (3)	— (0)
whoever	.1 (1)	— (0)
Total	2.2 (28)	— (0)

To- Infinitive Clause

The last two types of nominal subordinate clauses are of the nonfinite type. The fourth type described by Quirk et al. is the *to*- infinitive clause. Table 11 shows the frequency indices for *to*- infinitives in the spoken and written pear narratives (excluding the infinitives *to have* and *to be*). Following are some examples of the *to* construction found in the narratives.

EW2:14 He proceeds to confiscate one basket full of pears.

EW7:49 The cyclist gave the boy 3 pears to thank him & his friends.

ES2:147 He begins to notice that something's not right.

ES13:112 He starts to empty out his apron again.

Table 11 shows that *to*- infinitives are more frequent in the written than the spoken narratives. This is consistent with what Chafe (1982) found for *to*- complements in his sample, 2.7 in spoken discourse and 6.8 in written.

However, O'Donnell found more infinitives in spoken, 61 compared to 52 per 100 T-units for written. This disparity may be due to the greater formality involved in a television talk show (O'Donnell's sample) than in dinner table conversations (Chafe's sample) or in the retelling of the events of a movie (my sample).

TABLE 11
Frequency Indices for *to*- Infinitive Nominal
Subordinate Clauses in the Spoken and Written
Pear Narratives

	Spoken	Written
<i>to</i> - infinitive	2.5 (31)	7.6 (54)

TABLE 12

Frequency Indices for *-ing* Nominal Subordinate
Clauses in the Spoken and Written Pear
Narratives

	Spoken	Written
<i>-ing</i> clause	15.0 (189)	22.8 (161)

-Ing Clause

Nominal *-ing* clauses, or participial clauses, are the final type of nominal subordinate clause. These are also nonfinite clauses. (This does not include *-ing* clauses of verbal progressive constructions like 'He was picking pears', which is not a subordinate clause at all.) Some examples and the frequency indices for this construction in the pear stories appear below.

EW2:1 The film begins with a rather portly Mexican picking green pears.

EW12:29 Soon a little girl came riding from the other direction.

ES1:2 And there's a ladder coming out of the tree

ES11:69 So then we switch to the boy riding on the bicycle

These nominal clauses, like *to*- infinitives, are more frequent in writing: 7.8 more per 1000 words. Chafe's findings for present participle constructions are also in agreement, showing more present participle clauses in writing (15.8) than in speech (6.2). O'Donnell's findings are again in disagreement, showing more participle clauses in spoken, 11 per 100 T-units versus 9 in the written.⁶

Summary

Table 13 is a summary of all five types of nominal subordinate clauses analyzed in the spoken and written stories grouped into two basic types: finite and nonfinite. The finite clauses are *that*- complementizers, *wh*- interrogatives, and nominal relatives. The nonfinite clauses are *to*- infinitives and *-ing* clauses.

Table 13 shows that finite nominal subordinate clauses occur significantly more often in spoken narrative than in written. This is comparable to what O'Donnell found in his sample, i.e. 57 in spoken and 53 in written per 100 T-units. The counterpart to this abundance of finite nominals in speech is the

⁶Perhaps O'Donnell used different criteria for counting *to*- infinitives and *-ing* clauses. Unfortunately, he does not state his criteria nor give examples.

TABLE 13
Frequency Indices for All Nominal Subordinate Clauses in the
Spoken and Written Pear Narratives Grouped into two Types:
Finite and Nonfinite (Tables 8–12)

	Spoken	Written
Finite Clauses	10.9 (137)	3.8 (27)
Nonfinite Clauses	17.5 (220)	30.4 (215)
Total	28.3 (357)	34.2 (242)

greater frequency of nominalizations in writing (e.g. 'the man who was picking pears' versus 'the pear-picker'). Though not analyzed in the present study, Chafe's (1982) findings show 34.4 nominalizations in written discourse to 4.6 in spoken.

Though finite subordinate clauses are more common in the spoken stories, Table 13 shows that nonfinite clauses are more common in the written. Because nonfinite clauses pack more information into fewer words (e.g. 'the man picking pears' versus 'the man who was picking pears'), this distribution may be attributed to what Halliday calls the greater lexical density of writing.

Table 13 nicely supports Halliday's hypothesis that the structural complexities in each modality are of different natures: spoken narrative shows more finite subordinate clauses (low lexical density), whereas written exhibits more non-finite subordinate clauses (high lexical density).

Adjectivals

Relative clauses are the most common type of adjectival subordinate clause. The following discussion will present a description of the frequency, functions, and semantics of adjectival relative clauses in spoken and written narrative discourse.

For this analysis a broad, surface structure definition of the relative clause will be used. Relative clauses are adjectival subordinate clauses, beginning with a *wh*-pronoun, *that*, or $\emptyset$ (i.e. contact relatives, those in which the relative pronoun is deleted). This is analogous to Keenan's (1975) definition used in his work on relative clauses and universal grammar. Some examples of relative clauses from the pear stories follow.

ES3:21 He looks at a girl that was coming the other way

EW7:45 The boy who had been playing paddle ball

ES8:14 And there's a basket ($\emptyset$) he puts them in

TABLE 14
Frequency Indices and Percentages of Occurrence of each
Adjectival Relative Pronoun in the Spoken and Written
Pear Narratives

	Spoken	Written
Total Frequency Index	11.7 (147)	6.9 (49)
Restrictive Relatives	8.0 (101)	4.8 (34)
Nonrestrictive Relatives	3.7 (46)	2.1 (15)
who	33% (48)	31% (15)
which	11% (17)	29% (14)
that	46% (68)	20% (10)
($\emptyset$)	5% (7)	16% (8)
where	5% (7)	2% (1)
whom	— (0)	2% (1)
Total	100% (147)	100% (49)

Frequency Analysis

Table 14 shows the frequency index for relative clauses as well as the percentages of overall occurrence for each relativizer in the spoken and written pear stories.

Again, if we assume that syntactic complexity arises out of a greater use of subordinate clauses, a tendency that is characteristic of written language, our expectations are disconfirmed. There is a greater frequency of relative clauses in the spoken narratives than in the written, five more per every 1000 words. The hypothesis that written language contains more dependent clauses than spoken is clearly not true with regard to relative clauses in the present set of narratives.

Restrictive relative clauses are considerably more frequent in both the spoken and written stories than nonrestrictive relatives. A restrictive relative clause gives a necessary determination to the noun it modifies so as to keep it distinct from other things of that type with which it might be confused. For want of any other measure, the criterion used for distinguishing nonrestrictive relative clauses in the written stories is whether it is set off by commas (i.e. marked nonrestrictive by the subject; this of course does not allow for incorrect punctuation on the part of the subject); the criterion used in the spoken stories is whether the modified clause ends in clause-final intonation.⁷ Some examples follow.

⁷This procedure was suggested to me by Wallace Chafe, who recommended its use for lack of any other reliable and consistent measure.

Restrictive

EW20:68 the same path that the man and goat took

ES17:118 the little boy that was on the bike, had been wearing a hat

Nonrestrictive

EW20:33 He crosses paths with a girl, who is also on a bicycle

ES12:195 three boys, who now have polished and are eating their pears

Chafe (1982), considering only restrictive relative clauses, found an opposite distribution, 9.3 in his spoken sample and 17.1 in his written, 7.8 more relatives per 1000 words in the written. Kroll also found more relative clauses in her written sample and O'Donnell found 13 more adjectival clauses per 100 T-units in his written.

However, in a 1969 study looking at the differences between students' oral and written narratives, Golub (1969:83) found 'at the .05 level of significance, there are more relative clauses in high-oral than in high-written'. By 'high-oral' and 'high-written', Golub means those compositions rated 'high' by teachers versus those rated 'low'. He suggests that this greater abundance of relative clauses in spoken narrative 'may suggest that the growth of linguistic performance in oral discourse is somewhat in advance of written performance' (1969:84).

Four opposing hypotheses can be suggested as an explanation for this apparent disparity. First, since Chafe's written sample comes from academic papers, Kroll's from homework assignments, and O'Donnell's from published newspaper columns, it might be that relative clause formation is a feature of formal, planned, nonspontaneous discourse rather than of written discourse. Because the pear stories are relatively informal and unplanned narratives, this interpretation would explain the overall lower frequency of relatives in the present data (11.7 in comparison to Chafe's 17.1). Second, relative clause formation is not a complex construction in comparison to, say, participles, infinitives, or nominalizations. Third, the findings of previous studies are wrong, and spoken discourse, as Halliday suggests, is the more complex modality. Fourth, as Tanen suggests, narrative genre may be the key factor distinguishing these stories from other types of discourse.

It seems that all four of these hypotheses may be right in part: spoken language is more complex than has hitherto been supposed; the complexities involved in each modality are of very different types (e.g. spoken may use more relative clauses, whereas written may use more nominalizations). The differences that emerge between the two modalities are often reflections of the formality and purpose of the discourse rather than of a spoken/written distinction.

In reference to Table 14 again, a striking difference between the spoken and written usage of relative pronouns can be seen in the greater frequency of the relativizer *that* in the spoken narratives than in the written. That is, 46 percent of

the relative clauses in the spoken, or almost one half, use *that* rather than *who* or *which* as compared to only 20 percent in the written. By contrast we see a much greater use of *which* in the written narratives: 29 percent over 11 percent.[†]

The following examples from the spoken narratives show this preference for *that* to introduce a relative clause rather than *who* or *which*.

ES1:120 and one kid takes the rock that he tripped on

ES12:75 a young boy on a bicycle that was much too big for him

Comparable examples from the written narratives show the preference for *who* or *which*.

EW15:29 along comes three young boys who help him pick the pears up

EW13:24 and making him run into a rock which made him fall

Combining the values for *who*, *which*, and *whom*, spoken discourse shows a 44 percent usage, almost equal to the percentage of *that*, whereas written discourse shows 62 percent occurrence, far exceeding any of the other values. The spoken register, being in general more informal than the written, provides an explanation for the greater use of the more informal relativizer *that*.

It is interesting to note here that the pronouns *who*, *which*, and *whom* were originally only interrogative and came into usage as relative pronouns during the Middle English period. Previously, *that* had been the only relative pronoun. Jespersen (1933) contends that the *wh*-relatives reminded scholars of the corresponding Latin pronouns, hence their introduction into many learned texts at the expense of the more natural *that*. This more informal nature of the pronoun *that* explains its greater occurrence in the more casual spoken narratives and the more frequent occurrence of the more learned *who*, *which*, and *whom* in the written narratives.

Table 14 also shows a greater tendency for the relative pronoun to be deleted in written discourse than in spoken (16 percent versus 5 percent). Some examples follow.

EW8:15 stopping to wipe one with the kerchief (Ø) he had around his neck

ES17:149 they continue walking in the direction (Ø) they were walking in

Though Jespersen (1933:351) contends that the deletion of the relative pronoun is preferred in colloquial, informal speech 'as being more direct and straightforward', the data do not support this contention. Though the small number of occurrences in the data, 15, makes it hazardous to suggest any firm conclusions, a few preliminary remarks from work in the field of perceptual processing seem appropriate.

[†]Ed.'s note: Copy editors, however, do their best to reverse this trend. In this volume, for example, the copy editor changed many instances of *which* to *that*. This suggests the significant and unexamined (as far as I know) phenomenon that written discourse in print represents not the sole production of the author—as in the narratives analyzed here—but rather the more or less joint production of an author and one or more editors. This note is another case in point.

Fodor and Bever (1965) and others show that phrase or clause boundaries play an important role in the segmentation of speech for perceptual processing. Drawing on this research, Fodor and Garrett (1967) find that the presence of a relative pronoun makes the sentence perceptually less complex than the same sentence in which the pronoun has been deleted. This is presumably because the relative pronoun provides a 'surface structure clue to semantically crucial deep structure relationships' (p. 290). Such cues help the listener to segment sentences into their constituent parts for processing. On this measure, then, it seems that the written narratives exhibit a greater degree of complexity than the spoken in having a greater percentage of deleted relative pronouns.⁸

Functional Analysis

The foregoing analysis has considered only the form and not the function of the relative clause in the discourse. Though there are many types of functional analyses that could reveal aspects of syntactic complexity, the type that will be considered here is particularly enlightening as it evaluates the role of the relative pronoun with respect to the other elements in the clause in which it occurs. The Keenan-Comrie Case Hierarchy uses this technique to measure syntactic complexity between two different texts.

Keenan (1975) explains that the Case Hierarchy (CH) forms an implicational scale,

Subject $\geq$ Dir. Obj $\geq$ Indirect Obj $\geq$ Oblique $\geq$ Genitive $\geq$ Obj of Comparison

such that simple relatives occur on the subject, then the direct object, then the indirect object, and so on, with the most complex relatives being objects of comparison. (Keenan collapses the indirect object and oblique positions [i.e. objects of prepositions] in English because they behave the same way.)

Below are some examples from the pear stories of these different types of relative clauses.

Subject

ES10:173 and they ran across his hat that was blown off

EW12:44 spotted the hat that had fallen off

Direct Object

ES10:38 and put them in a couple of barrels that he's got down there

EW12:48 the owner of the hat gave him three pears for him & his friends, which he shared with them

⁸It is interesting to note here that deleted *that*- complementizers do not follow the same pattern in being perceptually more complex and hence more frequent in written discourse; rather the opposite appears to be the case (see Table 8).

Oblique

ES10:34 he has an apron that he puts them in

EW12:52 he looked puzzled and counted his baskets, of which there were only two.

The final two positions on the Case Hierarchy were not found in the pear stories; however, some hypothetical examples follow.

Genitive

The farmer whose pig John stole.

The gate the hinges of which were rusty.

Object of Comparison

The boy that Mary is shorter than.

Keenan provides data showing how authors who are judged to use syntactically simple sentences, such as George Orwell, relativize more often on the subject and progressively less often on the other positions on the CH than authors who are considered to use syntactically complex sentences, such as P. F. Strawson. An idealized scheme for typical complex and simple authors is presented in Figure 3.

Figure 3 shows that 'simple' authors relativize more frequently on the subject position and less frequently on the lower end of the scale than more 'complex' authors; 'complex' authors, however, show a much greater use of the

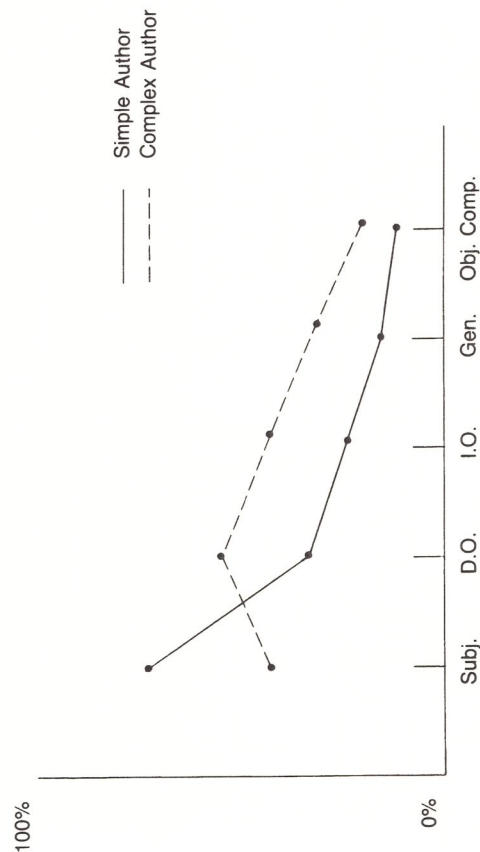


FIGURE 3. Idealized Predictions for the Typical Simple and Complex Authors' Use of Relative Pronouns with Respect to the Keenan-Comrie Case Hierarchy (based on Keenan's [1975:146] Figure 5)

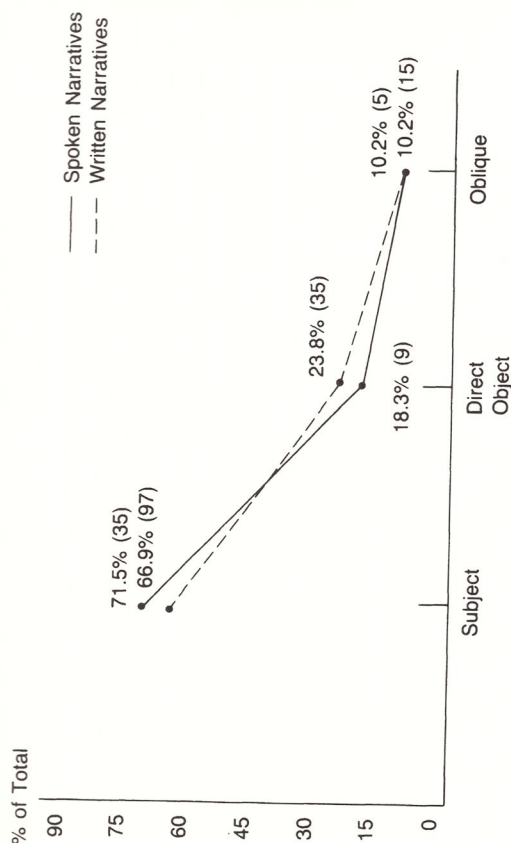


FIGURE 4. The Case Hierarchy for Adjectival Relative Clauses in the Spoken and Written Pear Narratives. Spoken: $N=147$; Written: $N=49$

more complex oblique relatives than 'simple' authors. This measure of syntactic complexity can be extended to determine whether similar differences exist between spoken and written narrative discourse.

Applying the Case Hierarchy analysis to the pear stories data, the hierarchy is confirmed in that the most relatives occur on the subject, then the direct object and then the oblique (Figure 4). However, there seems to be no difference in syntactic complexity between the spoken and written narratives as Keenan found between simple and complex authors. The percentages of occurrence of each type of relativizer are not significantly different between the spoken and written data. Thus, it seems that the use of relative clauses in spoken and written narrative discourse do not exhibit any functional differences on this measure of syntactic complexity.

Semantic Analysis

After considering the foregoing analysis of the form and function of relative clauses in discourse, it is important to look at the semantics, or the kind of information that is encoded in relative clauses in both the spoken and written narratives. One insight into the semantics of relative clauses is an analysis on the basis of the type of information that they contain.

Relative clauses can be subdivided into 'information-bearing' and 'identi-

ficator' types (Bernardo 1979); the former provide new information, whereas the latter specify old information. Based on Bernardo's distinction, some examples from the pear stories follow.

Information-Bearing

ES7:4 There was a man who was picking pears.

EW10:13 The three boys pass the man who had just come down the ladder.

Identificatory

ES3:47 And then the man who was picking pears comes down from his ladder.

EW3:49 The man who had been picking the pears is just getting down from the ladder.

The first examples are information-bearing relatives because they provide new information to the hearer about 'the man'. The second are identificatory relative clauses because they single out 'the man who was picking pears' from other members of the category 'man'. (This categorization can only be understood within the context of the entire narrative.)

Classifying the relative clauses in the pear stories on the basis of this distinction, the spoken and written narratives reveal the differences shown in Table 15. As the data show, relative clauses in both the spoken and written stories, for the most part, provide the reader/hearer with new information about the nouns they modify (i.e. 59 percent and 78 percent). What is of significance is the greater percentage of identificatory relative clauses in the spoken narratives, 19 percent more, than in the written. Relative clauses in the written narratives show a greater tendency to provide the reader with new information about the nouns they modify.

It seems that identificatory relative clauses are used more frequently in spoken discourse because of the speaker's need to ensure that the hearer is correctly perceiving the object or agent being talked about. This is largely a phenomenon of spoken language because the speaker may not be sure exactly

TABLE 15
Percentage of Occurrence for Information-Bearing and
Identificatory Relative Clauses in the Spoken and Written
Pear Narratives

	Spoken	Written
Information-bearing	59% (87)	78% (38)
Identificatory	41% (60)	22% (11)
Total	100% (147)	100% (49)

how much identificatory material is actually needed by the hearer; hence, the speaker monitors the communication channel in an effort to fulfill Grice's Cooperative Principle—the maxim of quantity: provide just as much information as is needed, no more.

This repetitious nature of speech (cf. Chafe 1982) is used to make up for the lack in speaking of precision and clarity more common to writing. In written discourse, not as much identificatory material is needed to enable the reader to understand who or what is being referred to because the very nature of the written mode provides greater clarity and precision (cf. Poole and Field 1976). In written discourse the information is merely stated and not elaborated; it is left to the reader to make of it what he or she will.

If we assume that the lack of abundant identificatory information and the introduction of considerable new information in subordinate relative clauses provides increased complexity, then the written narratives can be considered more complex on this measure. With reference to the previous discussion on perceptual processing, this assumption is confirmed if it is believed that it is perceptually more difficult to process new information occurring in a dependent clause than in a main clause.

Because of the greater number of identificatory relative clauses in the spoken stories, it is interesting to analyze what information is subordinated in the relative clauses and what types of constructions are used in the written narratives to convey the same information.

The largest number of relative clauses in the spoken narratives, 12 percent (16), refer to 'the man who was picking pears.' Only 4 percent of the relatives in the written stories convey this type of information. In the written narratives this information is generally revealed through the postnominal adjective, 'the man picking pears' (cf. Table 12 showing greater use of nominal *-ing* subordinate clauses in the written stories), or nominalizations such as 'the pear-picker', 'the picker', 'the fruit-picker', or simple nouns such as 'the worker', 'the farmer', 'the laborer', or simply 'the man' (cf. Chafe's nominalizations in written).

Other information frequently conveyed in relative clauses in the spoken stories are of the following types: 'the baskets that he puts the pears in', 'the boy who returned the hat', 'the man who is on the ladder', and 'the boy who is on the bike'. All of these relatives are identificatory-type relative clauses providing clarity to the listener about the object or agent being referred to.

This distinction in the use of relative clauses in spoken and written language can be explained with reference to Halliday's concept of lexical density. As relative clauses show a decrease in lexical density, we would expect to find more in the spoken narratives. At the same time, we expect to see more nominalizations (as Chafe did) and more nonfinite subordinate *-ing* clauses (as Table 12 shows), showing an increase in lexical density in written discourse as the counterpart.

Adverbials

The third type of subordinate clause that will be considered in this analysis is the adverbial subordinate clause. Some examples follow.

EW8:5 When the apron was full, he climbed carefully down

EW1:19 As they passed, he turned to look at her

EW1:4 While he was doing this another man walked by leading a goat

EW1:9 and he doesn't seem to be paying all that much attention because you know the pears fall

ES10:25 I wasn't sure at first if they were apples

Table 16 shows the frequency indices for all adverbial subordinate clauses in both the spoken and written pear stories. Though the figures are small for the individual subordinators, the overall totals are significant. Overall, adverbial subordinate clauses are considerably more frequent in the written narratives (13.0) than in the spoken (8.4). This is in agreement with what O'Donnell found for adverbial clauses in his sample, i.e. 22 for spoken and 33 for written per 100 T-units.

TABLE 16
Frequency Indices for Adverbial Subordinate Clauses in the
Spoken and Written Pear Narratives

	Spoken	Written
when	1.6 (20)	3.7 (26)
as	1.3 (16)	3.7 (26)
while	.8 (10)	2.3 (16)
because	1.3 (16)	— (0)
if	1.3 (17)	— (0)
whether	— (0)	.8 (6)
after	.2 (3)	.6 (4)
where	.5 (6)	.4 (3)
like	.9 (11)	.3 (2)
since	.2 (2)	.3 (2)
(al)though	.1 (1)	.1 (1)
as if	— (0)	.3 (2)
before	.2 (2)	.1 (1)
so that	.2 (2)	— (0)
wherever	— (0)	.1 (1)
whereby	— (0)	.1 (1)
once	— (0)	.1 (1)
Total	8.4(106)	13.0 (92)

TABLE 17
Semantic Distribution of Adverbial Subordinate Clauses in the
Spoken and Written Pear Narratives

	Spoken	Written
Time	50% (53)	83% (76)
Condition	17% (18)	10% (9)
Reason	15% (16)	0% (0)
Place	6% (6)	4% (4)
Other	12% (13)	3% (3)
Total	100% (106)	100% (92)

Adverbial clauses can be grouped into different types based on their semantics, i.e. time, condition, reason, place (Quirk et al. 1972:774-745), as Table 17 shows. Time adverbials are those with *when*, *as*, *while*, *after*, *since*, and *before* subordinators. Condition adverbials, which also include concession and hypothetical adverbial clauses, have *if*, *as if*, *whether*, and *(al)though* subordinators. Reason or cause adverbials are those with the subordinator *because*, and place adverbials are *where* or *wherever*.

Time adverbials are by far the most common type of adverbial subordinate clause.⁹ This is probably due to the fact that the data are narratives in which subjects were retelling a sequence of events. It is interesting to see a significantly greater percentage of time adverbials in the written narratives—33 percent more. This is presumably because writing has the characteristic of what Halliday calls 'not anchored in the here and now', what Chafe calls 'detached', and what Lakoff calls minus 'visibility'. Because of this, more time adverbials are needed by the writer to clarify the sequence of events to the reader, whereas extralinguistic factors are available to the speaker to provide this information.

In spoken narrative, the common *and then* structure is more frequent for detailing the sequence of events. The greater overall use of 'time adjuncts' (Quirk et al. 1972:482) in spoken narrative can be seen in Table 18.

Only the *then* time adjunct is more frequent in the written stories. *Then* with a preceding coordinator (e.g. *and then*, *but then*, *so then*) is a more common structure in the spoken narratives, 9.3 (117) for the three constructs together, than in the written, 1.3 (9). This provides further evidence that instances of the coordinator *and* in spoken discourse are fulfilling functions other than that of pure coordination. A large percentage of the *and*'s in the spoken narratives function analogously to subordinate time adverbials or time adjuncts, as seen in

⁹Unfortunately Kroll's figures are not comparable with those of the present study because she did not consider the functional differences (i.e. nominal, adjectival, adverbial) in subordinate clauses.

TABLE 18
Frequency Indices for Time Adjuncts in the Spoken and Written
Pear Narratives

	Spoken	Written
and then	8.8 (111)	.9 (7)
then	2.1 (26)	5.8 (41)
and so	1.3 (16)	— (0)
but then	.3 (4)	.2 (2)
so then	.1 (2)	— (0)
Total	12.6 (159)	7.0 (50)

the following examples from the episode where the young boy first appears in the film.

Spoken

- ES1:47 And then definitely when he's up there, a kid comes by on a bicycle.
 ES3:11 And then this kid comes along with a bicycle.
 ES5:18 And then this little boy is riding by on his bicycle.
 ES8:18 And then a boy comes by on a bicycle.
 ES9:20 And he's up there picking, and a little boy comes by on his bicycle.
 ES12:42 And then a little kid happened by on a bicycle.

Written—Time Adverbial Subordinate Clauses

- EW5:17 As he was picking more, a small boy on a relatively big bicycle came by & stole 1 whole basket of fruit,
 EW7:10 While the man was up in the tree a boy on a bike came by & stopped.
 EW10:5 While he was up picking pears, a boy came riding up on his bicycle.

Written—Other Time Expressions

- EW1:6 Then a boy rode up on a bicycle.
 EW2:12 Next on the scene is a boy in a large-brimmed hat riding a red bicycle.
 EW8:21 Then this kid comes ridding by on this bicycle.

Table 17 also shows that, as opposed to the greater frequency of time adverbials in the written narratives, the spoken stories show a greater percentage of both condition and reason adverbial subordinate clauses. It is interesting to note that there are no *because* (reason) or *if* (condition) clauses in the written stories; however both occur with a 1.3 frequency in the spoken stories (Table 16). This may be because speakers feel a greater need to justify and give explicit reasons for their claims, whereas writers, who may feel the need, do so with devices other than subordination.

It should be noted here that Chafe (1982) considers *because* a coordinator,

as its distribution is closer to that of coordinators (i.e. greater frequency in speech) than subordinators. However, since the semantic function of *because* clauses is closer to that of subordinators in that the information is presupposed rather than asserted (cf. discussion above), it is more appropriately classed as an adverbial subordinator.

Summary

In summary, then, Table 19 presents the total frequency indices for all three types of subordinate clauses: nominal, adjectival, and adverbial. Directly contrary to many previous assumptions, subordinate clauses are frequent in spoken as well as in written narrative. The nominal subordinate clause is, overall, the most frequent type of subordinate clause in both modalities. Whereas nonfinite nominals and adverbial subordinate clauses are more frequent in written narrative, finite nominals and adjectival subordinate clauses are more common in spoken. Thus, we see differing types of complexity in each modality.

CONCLUSION

This study has presented an in-depth analysis of subordination and coordination in narrative discourse. Based on the assumption that subordination implies complexity, the results show that, contrary to many previous assumptions, spoken narrative is on the whole just as complex as, if not more complex in some respects, than written narrative. Most significant, perhaps, is that, as Halliday suggests, the types of complexities involved in the two modalities are different. The increased lexical density of writing is balanced by a relatively greater number of dependent clauses in speech. This 'integrated' (Chafe) and 'compact' (O'Donnell) nature of written discourse may be largely a result of the increased 'planning time' (Ochs) allowed to the writer and of the lack of 'visibility' (Lakoff) and absence of extralinguistic factors between writer and audience.

TABLE 19
Frequency Indices for All Subordinate Clauses in the Spoken and Written Pear Narratives

	Spoken	Written
Nominals: Finite	10.9 (137)	3.8 (27)
Nonfinite	17.5 (220)	30.4 (215)
Adjectivals	11.7 (147)	6.9 (49)
Adverbials	8.4 (106)	13.0 (92)
Total	48.4 (610)	54.2 (383)

This study has also shown that previous studies have often confounded the spoken/written distinction with differing degrees of formality and purpose of the discourse. Though it is difficult to obtain two samples of discourse that are similar in all respects except modality, Lakoff's criteria for evaluating the position of the text on the discourse continuum provide a valuable means for finding comparable discourse samples that differ by only the presence or absence of a single feature. Narrative discourse, elicited under controlled conditions, proves to be one such discourse type.

Because of this problem, previous studies have found conflicting and opposing results regarding the type and degree of complexity found in spoken and written language. Differences in syntactic complexity between the spoken and written modalities which previous studies have found often turn out to result from differences in the formality and purpose or register of the discourse rather than true differences between spoken and written language.

Thus it seems that many forces are at work, pushing and pulling discourse in different directions. Though written language should exhibit greater complexity because of a lesser degree of spontaneity, a relatively greater degree of formality, and the increased planning time afforded the writer, militating against this is the lack of visibility and extralinguistic factors to disambiguate the written text and the emphasis on clarity and simplicity of style by writing teachers. In addition, the consequential and more permanent nature of most instances of writing causes writers to be more conservative in their expression than speakers need to be.

This study has also suggested that a large number of the *and* coordinators creating the 'fragmented' nature of speech, are fulfilling functions other than that of coordination and hence should not be treated as coordinators. Evidence for this has been suggested by the preponderance of *and* to initiate a sentence, serving as a filler word or weak connective, by its often reduced phonological expression as [ən] or [n], and by its coupling with the time adjunct *then* to relate the sequence of events in a narrative.

In addition it has been shown that merely evaluating the presence of dependent clauses in a discourse text is insufficient to account for complexity in language. A combined functional and formal classification of subordinate clauses reveals that though nonfinite nominals and adverbial subordinate clauses are more common in written narrative, finite nominals and adjectival relative subordinate clauses are more frequent in spoken narrative. Directly contrary to previous studies, it has been found that spoken narrative, as well as written, exhibits a high frequency of adjectival relative clauses. Though the spoken narratives show greater use of restrictive relative clauses, a larger percentage of the relative clauses in the written narratives are of the information-bearing type.

Overall, this study has provided some new insights into the problems and complexities involved in the analysis of spoken and written discourse. Perhaps most important, it has shown that the basic assumption of many linguists in the

past, i.e. that subordination implies complexity, is false. The evaluation of syntactic complexity is simply more complex than that.

REFERENCES

- BERNARDO, ROBERT. 1979. The function and content of relative clauses in spontaneous oral narratives. *Proceedings of the Fifth Annual Meeting of the Berkeley Linguistics Society*, 539–51.
- CHAFE, WALLACE. 1979. Integration and involvement in spoken and written language. Paper presented at the 2nd Congress of the International Association for Semiotic Studies, Vienna.
- CHAFE, WALLACE. 1980. The deployment of consciousness in the production of a narrative. *The pear stories*, ed. by Wallace Chafe, 9–50. Norwood, NJ: Ablex.
- CHAFE, WALLACE. 1982. Integration and involvement in speaking, writing, and oral literature. *Spoken and written language: Exploring orality and literacy*, ed. by Deborah Tannen, 35–53. Norwood, NJ: Ablex.
- FODOR, J. A. AND T. G. BEVER. 1965. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior* 4:414–20.
- FODOR, J. A. AND M. F. GARRETT. 1967. Some syntactic determinants of sentential complexity. *Perception and psychophysics* 2:289–96.
- GOLUB, LESTER S. 1969. Linguistic structures in students' oral and written discourse. *Research in the Teaching of English* 3:70–85.
- HALLIDAY, M. A. K. 1979. Differences between spoken and written language: Some implications for literacy teaching. *Communication through reading: Proceedings of the Fourth Australian Reading Conference*, ed. by Glenda Page, John Elkins, and Barrie O'Connor, 37–52. Adelaide, S. A.: Australian Reading Association, Vol. 2.
- JESPersen, OTTO. 1933. *Essentials of English grammar*. University, AL: University of Alabama Press.
- KEENAN, EDWARD L. 1975. Variation in universal grammar. *Analyzing variation in language*, ed. by Ralph Fasold and Roger Shuy, 136–48. Washington, DC: Georgetown University Press.
- KROLL, BARBARA. 1977. Combining ideas in written and spoken English: A look at subordination and coordination. *Discourse across time and space*, ed. by Elinor Keenan & Tina Bennett, 69–108. Southern California Occasional Papers in Linguistics, No. 5. Los Angeles, CA: University of Southern California.
- LABOV, WILLIAM. 1972. The transformation of experience in narrative syntax. *Language in the inner city*, 354–96. Philadelphia, PA: University of Pennsylvania Press.
- LAKOFF, ROBIN. 1979. Expository writing and the oral dyad as points on a communicative continuum: Writing anxiety as the result of mistranslation. Unpublished manuscript.
- MARTINET, ANDRÉ. 1964. *Elements of general linguistics*. London: Faber and Faber.
- OCHS, ELINOR. 1979. Planned and unplanned discourse. *Discourse and syntax*, ed. by Talmy Givón, 51–80. New York: Academic Press.
- O'DONNELL, ROY C. 1974. Syntactic difference between speech and writing. *American Speech* 49:102–10.
- POOLE, MILLICENT E. AND T. W. FIELD. 1976. A comparison of oral and written code elaboration. *Language and Speech* 19:305–11.
- QUIRK, RANDOLPH, SIDNEY GREENBAUM, GEOFFREY LEECH, AND JAN SVARTVIK. 1972. *A grammar of contemporary English*. London: Longman Group Limited.
- TANNEN, DEBORAH. 1980. Spoken/written language and the oral/literate continuum. *Proceedings of the Sixth Annual Meeting of the Berkeley Linguistics Society*, 207–18.
- TANNEN, DEBORAH. 1982. Oral and literate strategies in spoken and written narrative. *Language* 58:1. 1–21.