A taxonomy of revision changes was developed and applied to 18 case studies of writers' revisions. Subjects were six inexperienced student-writers, six advanced student-writers, and six expert adult-writers. The primary distinction of the taxonomy was between surface (formal and meaning-preserving) revisions and text-base (microstructure and macrostructure) revisions. Over a three-day span, the subjects (1) thought and made notes about the topic, (2) wrote an essay, and (3) wrote a revision/second draft. In addition, the expert writers made revisions of the inexperienced writers' first drafts. The results showed certain characteristic differences in how the different groups of writers revised their work. The experts turned out to be the most infrequent revisers. The inexperienced writers' changes were overwhelmingly surface changes, while the experts' changes were more evenly distributed between surface and text-base changes. The advanced students were the most frequent revisers of the three groups, making surface changes as often as did the inexperienced writers and text-base changes much like those of the experts. The experts' changes of the inexperienced writers' first drafts were predominantly macrostructure changes, particularly additions, substitutions, and distributions. The experts were much more uniform in these revisions than in their revisions of their own texts. (RL)
USING TEXT STRUCTURE MODELS FOR ANALYZING REVISION

“PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Lester Faigley

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Lester Faigley
Department of English
University of Texas at Austin
Austin, Texas 78712

Read at the Annual Meeting of the Conference on College
Composition and Communication, Dallas, March 1981

Stephen Witte was a coinvestigator for this project. The
research reported was supported by the National Institute of
Education, Department of Education, under grant number
NIE-G-80-0054. The opinions expressed herein do not necessarily
reflect the policy of the National Institute of Education.
Recent research in the composing process (Flower & Hayes, 1980; Sommers, 1980) and statements by writing teachers (Hairston, 1981; Murray, 1979) indicate that experienced writers often make structural and content changes in their work during revision. Scholars studying literary manuscripts have long noted that famous writers use complex revision strategies, but the few attempts to classify the types of revision changes in the genesis of literary texts have been impressionistic (e.g., Hildick, 1965). The National Assessment of Educational Progress (Rivas, 1977) used a similar system to gauge the effect of revision changes on texts in its 1977 survey, using categories such as organizational, stylistic, continuational, and holistic changes. Both the NAEP taxonomy and Hildick's system do not provide an adequate scheme for classifying revision changes. The most rigorous studies of students' revisions to date have been those of Sommers (1978, 1980) and Bridwell (1980). Sommers classified changes by length and by type of operation, using the same categories (deletion, addition, substitution, rearrangement) that Chorsky (1965) used to group transformations. Although Sommers' study advanced research in revision, measuring the effect of revision on the meaning of texts lies beyond the scope of her study. Bridwell employed a classification system similar to Sommers'. In addition, Bridwell included a category for broad, text-motivated changes, but she found no examples in her 12th grade writing sample. Even though some descriptive research (e.g., Sommers, 1978, 1980) has shown that experienced writers revise more frequently and change longer units of text than inexperienced writers, no study has been able to describe
satisfactorily either the nature of these changes or their effect on the meaning of a text. Classification schemes based on syntactic theory are simply inadequate for these tasks.

The present study explores how the revisions of expert and inexperienced writers affect meaning. Part I describes a simple, yet robust, system for analyzing the effect of revision changes on the semantic structure of a text. Part I briefly summarizes two studies of writers revising which use the taxonomy set out in Part I. It also discusses the implications of these investigations for revision research, suggests directions for further research, and identifies caveats implicit in the conceptual framework and methodology of the studies reported.

I. A TAXONOMY OF REVISION CHANGES

Ten years ago there was little research in text structure to draw on for explaining the effects of revisions. But that situation has changed. Research in several disciplines—primarily text linguistics, cognitive psychology, and artificial intelligence—now attempts to account for semantic relationships among elements in a text. As a result, several key properties of text structure have been isolated. Research in particular has examined text coherence, discovering that besides internal, explicit cohesive devices (Halliday & Hasan, 1976), nonexplicit text inferences must also be analyzed in order to understand coherence (Clark, 1977; Crothers, 1978, 1979; Schank, 1975). Another essential property is text hierarchy. Work in discourse comprehension demonstrates that readers stratify
information according to levels of importance, remembering what is essential to the "theme" or "gist" of a particular text (van Dijk, 1980; Mintsch, 1974; Mintsch & van Dijk, 1979; Meyer, 1975; Thorndyke, 1977).

Using this corpus of research, a classification system that can account for how revision changes affect text structure can begin to take shape. Such a system must differentiate changes which affect the meaning of a text and changes which do not. For example, the addition of a comma in sentence (2) does not change the meaning of sentence (1), even though placement of a comma after an introductory subordinate clause is considered standard usage:

1. Because the horse lost his shoe the rider was lost.
2. Because the horse lost his shoe, the rider was lost.

In other instances, the addition of a comma does change the meaning of a text. The difference in punctuating restrictive and nonrestrictive modifiers is a classic case in point:

3. The governors who took bribes abused the public trust.
4. The governors, who took bribes, abused the public trust.

Sentence (3) of course refers only to the governors who took bribes while sentence (4) refers to all governors.

The primary distinction in our taxonomy is between those revision changes which affect the meaning of the text and changes
which do not affect the meaning of the text. This distinction, of course, is not a simple one, but neither is it impossible to make. Meaning in texts has been represented as a TEXT BASE. Philosophers and linguists have described meaning with a formal language. The formal language most commonly used is some type of predicate calculus, but other, more complex languages such as Montague grammar (1974) have also been advanced. Psychologists have used a simplified predicate calculus to represent a TEXT BASE as a sequence of propositions containing a predicate or case relationship and one or more arguments (e.g., Frederiksen, 1975; Kintsch, 1974; Meyer, 1975). In the present system, we will use the notion of a TEXT BASE to refer to the meaning of a text.

Explicit TEXT BASES are typically incomplete; that is, speakers or writers rely upon the listeners' or readers' ability to make inferences. Consider the following short text with and without sentence (5A):

5. I just made it to the station on time.
5A. I got on the train.
6. I had to buy my ticket from the conductor.

Suppose sentence (5A) does not appear in the explicit text. If readers know that conductors sell tickets to boarded passengers who have not purchased tickets, then they understand implicitly that the narrator of this short text had to get on the train. Crothers' model would represent sentence (5A) as an implicit proposition if it were not stated explicitly, a procedure
consistent with interactive models of text processing used in artificial intelligence research (see Minsky, 1975; Schank, 1975; Schank & Abelson, 1977).

A completely explicit text is difficult to process, not only because such a text is verbose but because the reader or listener seeks to make additional connections which are not intended by the speaker or writer (cf. Shuy & Larkin, 1979). Sensitive to the inferential nature of texts, Crothers' (1979) model for representing a TEXT BASE accommodates implicit propositions and connectives as well as explicit propositions. Such a text-base model permits us to distinguish between additions that make explicit what can be inferred and additions which bring new information to the text.

The basic distinction between the left and right branches in Figure 1 below is whether new information is brought to the text or whether old information is removed in a way that it cannot be recovered. Stated most simply, TEXT-BASE changes alter the meaning of a text, NON-TEXT-BASE or SURFACE CHANGES do not. SURFACE CHANGES do not introduce new information, nor do these changes delete information that is not recoverable by inferencing. If information which readers can recover from a text through inferencing is made explicit through revision, the change will not affect the TEXT BASE. TEXT-BASE CHANGES, on the other hand, introduce, delete, or alter information that cannot be inferred or recovered.

INSERT FIGURE 1 ABOUT HERE
SURFACE CHANGES

SURFACE CHANGES are detailed on the left branch of Figure 1. The first node under SURFACE CHANGES is FORMAL CHANGES. These changes include a variety of operations associated with copy-editing such as most (but not all) spelling changes, changes in tense, number, and modality, expansions of abbreviations to their full forms, internal and endstop punctuation changes, and format changes. Capitalizations are included under spelling changes. Format changes are divided into two subcategories: paragraphing and an "Other" category, which includes changes such as blocking a long quotation. Also included under FORMAL CHANGES are changes conditioned by morphology and syntax. These changes include changes in case, tense, number, and modality.

The second major subcategory of SURFACE CHANGES includes MEANING-PRESERVING CHANGES, changes which "paraphrase" the concepts in the text base but do not alter them. ADDITIONS are the first type of change. MEANING-PRESERVING ADDITIONS raise to the surface what can be inferred. Sentence (5A) in the train example is a MEANING-PRESERVING ADDITION. The second type of MEANING-PRESERVING CHANGE, DELETIONS, represent the opposite process, where a reader is forced to infer what had been explicit. These two types of meaning-preserving changes commonly involve the addition or deletion of single words, often adverbs such as next or then.

MEANING-PRESERVING SUBSTITUTIONS trade words that represent the same concept. The changing of car to vehicle is one example, if
the referent is the same for both words. PERMUTATIONS involve rearrangements or rearrangements with substitutions. DISTRIBUTIONS occur when material in one text segment is passed into more than one segment. A change where a writer revises what has been compressed into a single unit into more than one unit is a distributional change. CONSOLIDATIONS represent the opposite process. In consolidations, material in two or more units is collected into one unit.

TEXT-BASE CHANGES

A second important distinction is introduced under TEXT-BASE CHANGES. Here we try to separate revision changes that are simple adjustments or elaborations of existing text and those changes that would alter the summary of that text. Research in how readers comprehend texts has recently sought to understand how meaning in a text is processed as it is read and how a gist of that text—a global notion of what the text is about—is constructed in memory. Kintsch and van Dijk (1978), Kintsch and Vipond (1979), and Vipond (1980), have tested a model that describes how readers process text at local and global levels. This model represents meaning at two levels: a MICROSTRUCTURE level where all ideas in the text are represented (including those ideas that can be inferred), and at a summary or MACROSTRUCTURE level that is abstracted from the microstructure. The conceptual representation of a summary with a MACROSTRUCTURE comes from the work of van Dijk (1972, 1977a, 1977b, 1980).

A summary, of course, can vary widely in comprehensiveness.
Books are sometimes published in condensed versions which reduce the original length by about half; study guides for particular books reduce original texts still further; encyclopedias and other reference works contain one-page summaries of books; and at the extreme, best seller lists often summarize books with a single phrase. For our purposes, a MACROSTRUCTURE is synonymous with a conceptual summary at the first level above the existing text. Thus in the present study, a MACROSTRUCTURE more closely resembles a condensation than an abstract.

MACROSTRUCTURES are derived from a text base by three types of operations (see van Dijk, 1980, pp. 16-50). The first and most general operation is deletion, which in terms of processing is a selection rule, determining which propositions will be retained in long-term memory. Propositions which are not necessary for the interpretation of another proposition, or propositions which have local relevance only, are not represented in the MACROSTRUCTURE.

A second operation is GENERALIZATION, in which a series of propositions are grouped into a macroproposition. A series of examples about a global topic can be subsumed into a single proposition at a higher level of abstraction:

(7) Mary was wearing an orange print skirt. Linda had on a purple pants suit. Cindy wore a red jumper.

According to van Dijk, such sequences neither are conditionally linked nor express stereotypical actions or settings. Yet from
such details we can construct a more abstract proposition, The women wore colorful clothing.

A third operation is CONSTRUCTION, in which a joint series of propositions defines a macroproposition. The constructed macroproposition denotes conventional knowledge about the world, described as scripts or frames. From a sequence such as (A),

(A) I walked out of the house and got in the car. It started rough, and I let it idle for a few minutes. I was late enough so that most of the traffic had cleared off of the freeway. I pulled in the parking garage under my building at 8:45.

we could construct the macroproposition, I drove to work.

Finally, in some cases the microlevel of a text is the same as the macrolevel. In very short texts, such as traffic warning signs, the microstructure and macrostructure often coincide. In longer discourses, headings and topic sentences may pass directly to the macrolevel.

MACROSTRUCTURE CHANGES, accordingly, are changes in the text base that would alter a summary of that text. In contrast, MICROSTRUCTURE CHANGES affect meaning but do not affect the summary of a text. A MICROSTRUCTURE ADDITION, for example, can be thought of as an extrapolation of the existing TEXT BASE. The operations under MICRO- and MACROSTRUCTURE CHANGES are the same six described under MEANING-PRESERVING CHANGES. In the case of both MICRO- and MACROSTRUCTURE changes, however, the text base is altered.
A computational analysis of revision changes using the taxonomy described in Figure 1 is facilitated if the subcategories in Figure 1 are coded numerically. Each terminal node or "leaf" in the principal branches of the taxonomy has been numbered. Table 1 lists the subcategories from Figure 1 and their numerical codes.

Table 1

One other dimension of revision changes needs to be addressed—the span of text involved in the change. Recent research studies of revision have all examined the length of each change. We see the Bridwell (1980) system as having certain advantages over other systems. Our text-span taxonomy follows Bridwell's (1980) text-span classes with small exceptions. Our six classes are represented in Table 2.

Table 2

The text-span taxonomy is based on surface features for ease of use. If the number of propositions involved in a given change were computed, there would be little hope of using our taxonomy for comparing a number of texts. Furthermore, representations of propositions often coincide with clause and sentence boundaries. CLAUSES in this taxonomy are defined as constructions with finite subjects and verbs. Constructions longer than one word without both a finite subject and verb are classified as PHRASES. SENTENCES are determined by the writer's punctuation. If a
phrase or subordinate clause is punctuated as a sentence, it is counted as a sentence. MULTISENTENCE changes are classified one sentence at a time. Thus each sentence in a long MACROSTRUCTURE ADDITION would be labelled with a "6" for a multisentence addition and with a "31" for the effect upon text structure.

II. TWO STUDIES OF WRITERS REVISI NG

Eighteen cases studies of writers revising were collected and analyzed using the model described in Part I. Subjects were divided into three groups: inexperienced student writers, advanced student writers, and expert adult writers. Each student writer was enrolled in a writing class at the University of Texas. The six inexperienced students were recruited from a writing laboratory designed for students deficient in writing skills. The six advanced students were recruited from an elective, upper-division expository writing class, which typically attracts able and motivated students. Expert adults were recruited from professional writers living in Austin. All had journalistic experience; three had published fiction.

The procedures for gathering data were similar to those used by Bridwell (1980). Subjects were tested over a three-day span. On the first day, subjects were presented a writing topic which asked them to describe a place in Austin that an out-of-town audience likely would not know about. The writing situation specified that the description would be published in an out-of-town newspaper. Subjects were asked to think about the topic and to make notes if they wished. On day 2, writers wrote
an essay on the topic they had been given the previous day. These writing samples were collected and photocopied, and the changes that the writers had made were analyzed as Stage 1 revisions. On day 3, the original first drafts were returned to the writers who then wrote a second draft. When they had finished, both drafts were collected and analyzed. All changes between the first and second drafts were analyzed as Stage 2 revisions. In process revisions on the second draft were analyzed as Stage 3 revisions.

The results show certain characteristic differences in the way the different groups of writers revised their work. Somewhat surprisingly, the expert adult writers turned out not to be most frequent revisers. In an analysis of the combined revision changes, the expert writers made some type of change on the average of 137 times per 1000 words in the final draft, while the inexperienced students made a change 173 times per 1000 words and the advanced students 236 times per 1000 words. More revealing were the kinds of changes the different groups made. Combined changes for the major categories are summarized in Table 3.

The inexperienced writers' changes were overwhelmingly surface changes. Only 12% of the inexperienced writers' revisions were TEXT-BASE CHANGES. The expert adults' changes, in contrast, were more evenly distributed. About 35% of their changes were TEXT-BASE CHANGES. The advanced students were the most frequent revisers of the three groups. The advanced students made SURFACE
CHANGES about as often as did the inexperienced writers and about twice as often as the expert adults. In TEXT-BASE CHANGES the advanced students revised more like the expert adults than the inexperienced students, especially in MACROSTRUCTURE CHANGES, where the frequencies of advanced students and expert adults are nearly alike.

These data point to an important fact to keep in mind when comparing the revisions of expert adults to inexperienced writers—that the first draft of an expert writer is often closer to the writer's intended result than the first draft of an inexperienced writer. Expert writers, by and large, do not need to make as many surface changes as inexperienced writers. As Sommers (1980) and Perl (1979) have indicated, writers excessively concerned with surface changes are unaware of the more extensive changes possible in revision. Inexperienced writers in the present study just as in other studies were largely bound to what they wrote in the first draft. Their second drafts showed very few major alterations.

Another crucial factor we found in our analyses of expert writers' revisions was the extreme diversity in the ways expert writers revise. For example, one expert writer in the present study made almost no revisions, another started with an almost stream-of-consciousness text which she then converted to an organized essay in the second draft, another limited his major revisions to a single long insert that made his original text half again as long, and another revised mostly by pruning. To supplement these protocols, we conducted a series of interviews
with practicing writers of various sorts, writers of fiction, journalists, technical writers, and academicians in several disciplines. Again we found considerable variation. Probably the most important thing we learned was that many writers do not revise much at all, especially if the writing task is a familiar one. Indeed, the popular view of skillful writers as extensive revisers may be to a large extent fostered by interviews with novelists and studies of literary manuscripts. Reporters, for example, often cannot revise because of deadlines, and other writers on the job described revision as a luxury they could seldom afford. The amount and types of revision changes are dependent upon a number of variables besides the skill of the writer. The nature of the task, the writer's familiarity with the task, the length of the task, the audience for which the writing is intended, and personal habits of composing all influence how an expert writer revises.

Case studies of expert writers revising may be of limited value to teachers who want to help their students to write better. The ability of some expert writers to perform revisions operations mentally before committing a text to paper make their first drafts and subsequent revisions hardly comparable to those of student writers. In order to gain a better idea of what revision strategies to teach inexperienced writers, we made copies of the first drafts of three of the inexperienced writers and presented them to the expert adults. We asked the skilled writers to revise the three student essays as if they had written it as a first draft. We then analyzed these changes and compared the
results to how the inexperienced writers revised their own texts in the second draft.

Again there were large differences in the ways that the two groups of writers revised. The adult changes were predominantly MACROSTRUCTURE CHANGES. In particular, the adults used three processes—addition, consolidation, and distribution. They condensed what the students had written and then either elaborated or added information to support the points the students had wanted to make. The expert writers' revisions of the students' essays were much more uniform than the revisions of their own texts. The experts used two basic strategies: they condensed what the students had written and they expanded. What the expert writers did during revision that the inexperienced writers did not do was to determine the intention of the text, what the text was attempting to communicate.

The ability to formulate and to keep in mind a purpose and an audience for a text may be the central factor in explaining why experienced writers use recursive strategies during composing, looking back frequently at what they have written (cf., Flower & Hayes, 1980; Sommers, 1978, 1980). Inexperienced writers typically move through a writing task nonstop; experts typically reread and assess what they have done in light of what they want to do. Clearly, more research needs to be done on how writers of varying abilities compose, especially research on the importance of perceived goals and the facility to access long-term memory.
Research on revising may be considerably advanced through the use of computers that record revision changes, their sequence, and how long writers pause during composing and during revision.
### TABLE 1

A Classification of Revision Changes

#### I. SURFACE CHANGES

**A. FORMAL CHANGES**

1. SPELLING -- 01
2. TENSE, NUMBER, AND MODALITY -- 02
3. ABBREVIATION -- 03
4. PUNCTUATION -- 04
5. FORMAT
   a. PARAGRAPH -- 05
   b. OTHER -- 06

**B. MEANING-PRESERVING CHANGES**

1. ADDITIONS -- 11
2. DELETIONS -- 12
3. SUBSTITUTIONS -- 13
4. PERMUTATIONS -- 14
5. DISTRIBUTIONS -- 15
6. CONSOLIDATIONS -- 16

#### II. TEXT-BASE CHANGES

**A. MICROSTRUCTURE CHANGES**

1. ADDITIONS -- 21
2. DELETIONS -- 22
3. SUBSTITUTIONS -- 23
4. PERMUTATIONS -- 24
5. DISTRIBUTIONS -- 25
6. CONSOLIDATIONS -- 26

**B. MACROSTRUCTURE CHANGES**

1. ADDITIONS -- 31
2. DELETIONS -- 32
3. SUBSTITUTIONS -- 33
4. PERMUTATIONS -- 34
5. DISTRIBUTIONS -- 35
6. CONSOLIDATIONS -- 36
<table>
<thead>
<tr>
<th></th>
<th>Text-Span Classes for Revision Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GRAPHIC CHANGES</td>
</tr>
<tr>
<td>2</td>
<td>LEXICAL CHANGES</td>
</tr>
<tr>
<td>3</td>
<td>PHRASAL CHANGES</td>
</tr>
<tr>
<td>4</td>
<td>CLAUSAL CHANGES</td>
</tr>
<tr>
<td>5</td>
<td>SENTENCE CHANGES</td>
</tr>
<tr>
<td>6</td>
<td>MULTISENTENCE CHANGES</td>
</tr>
</tbody>
</table>
TABLE 3

Frequencies of Combined Revision Changes per 1000 Words in Final Drafts for Three Groups of Writers

<table>
<thead>
<tr>
<th></th>
<th>FORMAL CHANGES</th>
<th>MEANING-PRESERVING CHANGES</th>
<th>MICRO-STRUCTURE CHANGES</th>
<th>MACRO-STRUCTURE CHANGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXPERT ADULTS</td>
<td>20.9</td>
<td>68.8</td>
<td>27.5</td>
<td>19.6</td>
</tr>
<tr>
<td>ADVANCED STUDENTS</td>
<td>45.4</td>
<td>134.3</td>
<td>37.2</td>
<td>19.3</td>
</tr>
<tr>
<td>INEXPERIENCED STUDENTS</td>
<td>40.2</td>
<td>112.1</td>
<td>19.5</td>
<td>1.3</td>
</tr>
</tbody>
</table>
TABLE 4

Comparison of Combined Revision Changes per 1000 Words for Inexperienced Writers Revising Their Own Drafts and Expert Writers Revising the Same Drafts

<table>
<thead>
<tr>
<th></th>
<th>Formal Changes</th>
<th>Meaning-Preserving Changes</th>
<th>Micro-Structure Changes</th>
<th>Macro-Structure Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Inexperienced Students</td>
<td>11.6</td>
<td>43.8</td>
<td>24.1</td>
<td>0</td>
</tr>
<tr>
<td>6 Expert Writers Revising Student Texts</td>
<td>2.4</td>
<td>8.0</td>
<td>8.5</td>
<td>35.1</td>
</tr>
</tbody>
</table>
REFERENCES


FIGURE 1

A TAXONOMY OF REVISION CHANGES

Revision Changes

Surface Changes

Formal Changes

Spelling
Tense, Number, and Modality
Abbreviation
Punctuation
Format

Meaning-Preserving Changes

Additions
Deletions
Substitutions
Permutations
Distributions
Consolidations

Text-Base Changes

Microstructure Changes

Additions
Deletions
Substitutions
Permutations
Distributions
Consolidations

Macrostructure Changes

Additions
Deletions
Substitutions
Permutations
Distributions
Consolidations
asked that a place be selected in which an oil-soilet
system had not yet been installed. The writing clinic as
specified that the description would be published in an
out-of-town newspaper. Subjects were asked to think about the
scene and to make notes of they wished. On day 2, writers wrote

1.0
elegant. Only 12% of the experienced writers' revisions were

merely editorial. In contrast, 35% of the novice writers' changes were
textual changes. The advanced students made the most frequent

revisions of the three groups. The advanced students made SURFACE
organized essays. In the second draft, another limited his major revisions to a single long insert that made his original text half as long and another revised mostly by omitting. To supplement these protocols, we conducted a series of interviews.
strategies to reach inexperienced writers, we made copies of the
best parts of their work and presented it to the expert adults. We asked the skilled writers to revise the three student essays as if they had written it as a first draft. We then analyzed these changes and compared the
typically move through a writing task retrospectively: they typically report and describe what they have done in light of what they want to do. Clearly, more research needs to be done on how writers of varying abilities compose, especially research on the importance of perceived goals and the facility to access long-term memory.
