KELLOGG W. HUNT

Recent Measures in Syntactic Development

Any teacher of English can tell a fourth-grade theme from a twelfth-grade theme. Probably anyone in this room could make still finer distinctions: he could tell the average fourth-grade theme from the average eighth-grade theme. Just how would he detect the difference? For one thing he would rely on word choice. The vocabulary of the average eighth grader is measurably different from that of the average fourth grader. But also the teacher would feel that some of the sentence structures used by the eighth grader were too mature to be used by a fourth grader. Sentence structure, not vocabulary, is my subject for this paper.

The educational researcher respects the teacher's intuitive sense of maturity, but he wishes he knew how to measure it quantitatively, by counting something—if only he knew what to count. He knows of course that it takes centuries to build up a science. All during the Middle Ages the alchemists were poking away at the information which eventually led to modern scientific chemistry. It took centuries to establish the science. The science of measuring syntactic maturity is barely emerging from the stages of alchemy. It scarcely deserves to be called a science at all. But we do know a few things.

For the last thirty years we have known at least three things about the development of language structure. First, as children mature they tend to produce more words on any given subject. They have more to say. Second, as children mature, the sentences they use tend to be longer. Third, as children mature a larger proportion of their clauses are subordinate clauses.[1]

In the last two years it has been possible to add a few more measures, and I will come to them later. But first let me turn back to the statement about subordinate clauses and try to make clear its significance for the teaching program. It would be worse than useless for a fourth-grade teacher to say to her students, "Now if you will go back to your last paper and add more subordinate clauses to the main clauses, you will be writing like Miss Hill's wonderful sixth graders or Miss Summit's wonderful eighth graders instead of my own miserable fourth graders." Such an approach would be worse than useless. But the facts behind so useless a statement are not useless; they are useful if we know how to use them. Let us look at some fourth-grade writings. We find pairs of main clauses like this:

> There was a lady next door and the lady was a singer.

Now an older student would not be likely to repeat the noun *lady*. He might rewrite

Dr. Hunt is Professor of English at Florida State University in Tallahassee. This article is based on an address delivered to the National Conference on Research in English, Chicago, Illinois, February 19, 1966.

[1]Dorothea McCarthy, "Language Development in Children," *Manual of Child Psychology*, ed. Leonard Carmichael. New York: John Wiley & Sons, Inc., 1954.

the two clauses in any of several ways: one way would be to reduce the second main clause to a relative adjective clause.

There was a lady next door who was a singer.

Now instead of two main clauses and no subordinate clauses, we have one main clause and one subordinate clause.

Let me give you a few more examples of pairs of fourth-grade main clauses. In every instance one of the main clauses could have been reduced to a relative adjective clause.

Moby Dick was a very big whale. He lived in the sea. (who lived in the sea.)

His owner was a milkman. The milkman was very strict to the mother and babies. (who was very strict . . .)

Once upon a time I had a cat. This cat was a beautiful cat. It was also mean. (who was a beautiful cat.)

One day Nancy got a letter from her Uncle Joe. It was her great uncle. (who was her great uncle.)

I have a new bicycle. I like to ride it. (which I like to ride.)

We have a lot on Lake Talquin. This lot has a dock on it. (On Lake Talquin we have a lot which has a dock on it.)

Today we went to see a film. The film was about a white-headed whale. (which was about a white-headed whale.)

The jewel was in the drawer. It was red. (The jewel which was red . . .)

Beautiful Joe was a dog, he was born on a farm. (that was born on a farm.)

One colt was trembling. It was lying down on the hay. (One colt which was lying down . . .)

A convenient way to measure the frequency of subordinate clauses is to divide the total number of clauses, both subordinate and main, by the number of main clauses. I will call this the "subordinate clause index."[2] It is expressed as a decimal fraction. The index will always be 1 (for the main clause) plus whatever number of subordinate clauses are attached to it.

I find that average fourth-grade writers have a subordinate clause index of about 1.3; that is, they write a subordinate clause three-tenths as often as they write a main clause. Average eighth graders have an index of 1.4; they write a subordinate clause four-tenths as often as a main clause. Average twelfth graders have an index of 1.68. They write a subordinate clause about six-tenths as often as a main clause. If you jump now to the superior adult writers who produce articles for *Harpers* and *Atlantic* you find that they have an index of 1.78: they write about seven-tenths as many subordinate clauses as main clauses. However, some mature article writers have much higher indexes. One had a score of 2.36, indicating that his average main clause had one and a third subordinate clauses related to it.

The general trend of development is fairly clear: for fourth grade the score was 1.3, for eighth 1.4, for twelfth 1.6, and for superior adults 1.7.

It would be interesting to go back to the grades earlier than the fourth to see if the number of subordinate clauses is smaller back there. Fortunately Professors O'Donnell, Griffin, and Norris at Peabody have provided us with data within the last year.[3] Reporting on their results for speech alone, they find a general increase in number of subordinate clauses from kindergarten to the seventh grade where the study ended. These kindergarten students have an index of 1.16. Putting their figures and mine together, we see that the trend is clear. From the first public school grade to the last the number of subordinate

[2]The "subordination ratio" which has been used for thirty years has usually been figured in another way.

[3]Published as an NCTE research monograph entitled *A Transformational Analysis of the Language of Kindergarten and Elementary School Children* (in press).

clauses increases steadily for every grade.

This tendency has implications for teaching language. Without ever using the words "main clause" and "subordinate clause," the language arts teacher who sees pairs of main clauses like those I have mentioned can show her students another way of saying the same thing.

One further refining statement can be made about subordinate clauses and the index of their frequency. There are three common kinds of such clauses: noun, movable adverb, and adjective. The other kinds, such as clauses of comparison, are uncommon. Though the total of all three increases with maturity, not all three increase equally. Noun clauses in general are no index of maturity: the number of them is determined instead by the mode of discourse, the subject matter, all the way from the early grades to maturity. Movable adverb clauses do seem to increase with maturity in the very early grades, but the ceiling is reached early, and after the middle grades the frequency of them tells more about mode of discourse and subject matter than about maturity. But adjective clauses are different. From the earliest grades to the latest the number of them increases steadily, and among skilled adults the adjective clause is still more frequent than it is with students finishing high school. We see, then, that the subordinate clause index is a team which moves ahead, but it moves ahead because one member does almost all the work. The other two sometimes pull ahead but sometimes pull back too, depending on factors other than mental maturity.

But of course subordinating clauses is not all there is to syntactic development. In every pair of examples I have given so far, it would have been possible to reduce one of the clauses still further so that it is no longer a clause at all, but merely a word or phrase consolidated inside the other clause. In this fashion two clauses will become one clause. The one clause will now be one word or one phrase longer than it was before, but it will be shorter than the two clauses were together. By throwing away some of one clause we will gain in succinctness. The final expression will be tighter, less diffuse, more mature.

Let me illustrate now with the same examples, and then back up the examples with figures to indicate that older students do indeed more often make one longer clause out of two shorter ones.

A clause with a predicate adjective can all be thrown away except for the adjective.

Once upon a time I had a *beautiful, mean* cat.
The *red* jewel was in the drawer.

Eighth graders write more than 150 percent as many single-word adjectives before nouns as fourth graders do.

If a clause contains a prepositional phrase after a form of *be* you can throw away all but that prepositional phrase.

The jewel *in the drawer* was red.
Today we saw a film *about Moby Dick.*

Eighth graders use such prepositional phrases to modify nouns 170 percent as often, and twelfth graders 240 percent as often, as fourth graders do.

If the clause contains a *have* you can often put what follows the *have* into a genitive form and throw away the rest.

I like to ride *my* new bicycle.
Our lot on Lake Talquin has a dock on it.

Twelfth graders used 130 percent as many genitives as fourth graders do.

If a clause contains a predicate nominal, it can become an appositive, and the rest can be thrown away.

There was a lady next door, *a singer.*
His owner, *a milkman,* was very strict to the mother and babies.
One day Nancy got a letter from her *great uncle* Joe.

Eighth graders wrote a third more appositives than fourth graders.

Often clauses with non-finite verbs can all be thrown away except for the verbs, which now become modifiers of nouns.

> Beautiful Joe was a dog *born on a farm.*
> One *trembling* colt was lying down on the hay.
> One colt, *lying down on the hay,* was trembling.

Eighth graders wrote 160 percent and twelfth graders wrote 190 percent as many non-finite verb modifiers of nouns as fourth graders did.

I have used this set of examples twice now, to show two different things: first, how it is that older students reduce more of their clauses to subordinate clause status, attaching them to other main clauses; and secondly, how it is that the clauses they do write, whether subordinate or main, happen to have more words in them.[4] Those extra added words are not padding. They are all that is left out of useless whole clauses when the padding has been thrown away. From a six-word clause five words may be thrown away, with only one word salvaged. So adding that one more word to some other clause indicates a substantial gain. Though an increase of one word in clause length may not sound very impressive, a gain of five words or so in succinctness is indeed impressive. What was said in two clauses totalling twelve words is now said in one clause of seven words.

It is not as if some fourth-grade teacher had said "Add one word to each clause you have written." Instead it is as if she had said, "In this sentence you can throw away all but one word or one phrase. You can consolidate that word or phrase with this other expression into a larger, more comprehensively ogranized, unit of thought."

Substantial evidence is accumulating that as school children mature they do indeed learn to put their thoughts into longer and longer clauses. My own first research dealt with children of strictly average IQ, that is, with children having scores between 90 and 110. I worked first with three grades, fairly widely spaced: fourth grade, when students are just beginning to write with some degree of comfortableness; twelfth grade when the student of average IQ writes about as well as he ever will, perhaps; eighth grade, half way between the beginning and end of that public school period. The clauses written by these fourth graders were 6.6 words long. Clauses by eighth graders were 20 percent longer and clauses by twelfth graders were 30 percent longer. But development does not stop there. The writers of articles for *Harpers* and *Atlantic* write clauses about 175 percent as long as those written by fourth graders of average IQ. In fact, in clause length, the superior adult is farther ahead of the average twelfth grader than the average twelfth grader is ahead of his little brother back in the fourth grade.

If the evidence is as sound as it seems to be, then one ought to be able to predict on the basis of it. If this tendency to lengthen clauses is a general characteristic of linguistic development, then one might predict in several directions. He might predict that if growth is fairly steady after the fourth grade, then it probably is perceptible before the fourth grade too. And if growth occurs in writing, then it probably occurs in speech too. If one is going to measure

---

[4] A clause is here defined as one subject or one set of coordinate subjects with one finite verb or one finite set of coordinated verbs. Thus *I went home* is one clause, and so is *Jim and I went home and rode our bikes.*

The average clause length for any body of writing, however long or short, is simply the total number of words divided by the total number of clauses. For a sentence such as *She said he ought to try harder,* there are 7 words and 2 clauses, so the average clause length is 3.5 words for that body of writing.

the development which occurs earlier than the fourth grade, of course, it is speech, not writing, he must study.

Fortunately the Peabody study has provided us with some confirming evidence within the last few months. Notice the slight but steady increase as I read these figures. The clauses spoken by these kindergarten children are 6.1 words long. For first graders, 6.7 words. For second graders 7.1 words. Third graders 7.2. Fifth graders 7.5. Seventh graders 7.8. At every grade level there is an increase in the clause length of their speech. Clause length plots as a smooth rising curve, all the way to the maturity of *Harpers* and *Atlantic* articles.

One might predict in yet another direction. He might predict that children with superior IQ's will have matured more in language structure at even an early age. Since my results are not conclusive at this time, I am not sure whether, as early as the fourth grade, children with IQ scores above 130 write, on the average, slightly longer clauses and write a larger proportion of subordinate clauses than fourth graders with average IQ. By the time children of superior IQ reach the twelfth grade, however, their superiority in clause-length is unmistakable. They are almost as far ahead of average twelfth graders as average twelfth graders are ahead of fourth graders. In fact, in clause length, twelfth graders with IQ above 130 are closer to writers of *Harpers* and *Atlantic* articles than they are to twelfth graders of average IQ.

These longer clauses written by older students are not produced by combining just two clauses, but by combining four or six or eight. Superior adults can combine a dozen clauses into one, by the process already briefly suggested.

So, for a third time, I suggest that teachers who understand the findings of language development research may be able to apply those findings in the classroom. For years teachers have occasionally combined pairs of clauses as we were doing here a few minutes ago. But so far as I know it has never occurred to anyone to show that six or eight or a dozen are often consolidated into one mature clause. It is by this process that little sentences grow into big ones.[5]

Here is a clause written by an average eighth grader. "He was a rare white whale with a crooked jaw." That consolidates five clauses. 1) He was a whale. 2) The whale was white. 3) The whale was rare. 4) The whale had a jaw. 5) The jaw was crooked. Average fourth graders do not ordinarily write like that. In fact, in five thousand clauses written by fourth graders we found a single nominal that resulted from as many as five of these consolidations only three times. Five is simply too many for a fourth grader, but he often consolidates three.

Despite this eighth grader's relative maturity, even he failed to consolidate clauses where he might have. He missed opportunities. He wrote:

> Moby Dick was a dangerous whale. People had never been able to catch him. He was a rare white whale with a crooked jaw. He was a killer too. He was long and strong.

There are many ways to consolidate this further and I won't rewrite the whole passage. The first two clauses could well be consolidated and so could the last two:

> Moby Dick was a dangerous whale that people had never been able to catch.
> He was a killer too, long and strong.

I am recommending, then, that throughout the elementary and secondary grades the process of clause-consolidation is one

[5] See "How Little Sentences Grow into Big Ones," a paper by Kellogg W. Hunt read at the NCTE's Spring Institute on New Directions in Elementary English, Chicago, March 7, 1966, to be published with the proceedings of that institute.

of the things which the language arts program should study. Transformational grammarians speak of this process as the result of embedding and deletion transformations.

Finally I want to describe to you a new unit of measurement which is very convenient for syntactic development research. It is certainly more significant than sentence length which is still reported to be the most widely used measure of language maturity.[6] To introduce this unit let me read a theme as written by one of our fourth graders. The theme is one sentence long.

> I like the movie we saw about Moby Dick the white whale the captain said if you can kill the white whale Moby Dick I will give this gold to the one that can do it and it is worth sixteen dollars they tried and tried but while they were trying they killed a whale and used the oil for the lamps they almost caught the white whale.

In sentence length this fourth grader is superior to the average writer in *Harpers* and *Atlantic*. Now let me cut that sentence up into the new units. Each unit will consist of exactly one main clause plus whatever subordinate clauses happen to be attached to or embedded within it.

1. I like the movie we saw about Moby Dick, the white whale.
2. The captain said if you can kill the white whale, Moby Dick, I will give this gold to the one that can do it.
3. And it is worth sixteen dollars.
4. They tried and tried.
5. But while they were trying they killed a whale and used the oil for the lamps.
6. They almost caught the white whale.

For lack of a better name I call these units "minimal terminable units." They are "terminable" in the sense that it is grammatically acceptable to terminate each one with a capital letter at the beginning and a period or question mark at the end. They are "minimal" in the sense that they are the shortest units into which a piece of discourse can be cut without leaving any sentence fragments as residue. They are thus "minimal terminable units." I wish I could call these units "the shortest allowable sentences" but instead I call them "T-units," for short. To repeat, each is exactly one main clause plus whatever subordinate clauses are attached to that main clause.

In ordinary prose about half the sentences consist of just one such T-unit. The other half of the sentences consist of two or more T-units, often joined with *and's*. Such sentences are "compound," or "compound-complex." Cutting a passage into T-units cuts each compound sentence or compound-complex sentence into two or more T-units. Now if it were true that as writers mature they put more and more T-units into their sentences, then sentence length would be a better measure of maturity than T-unit length. But such is not the case. Occasionally a very young student will string one T-unit after another after another, with *ands* between or nothing between. The passage I read a moment ago combined six T-units into one sentence. The result of this tendency is that my fourth graders average more T-units per sentence than superior adults do. That fact upsets sentence length as an index of maturity. That same fact explains why T-unit length is a better index of maturity than sentence length.

A useful name for the average number of T-units per sentence might be "main clause coordination index." It probably should not be called "sentence coordination index."

Now let us pull all these various indexes together into a single piece of arithmetic. "Average clause length" is the number of

[6] For instance, see the article on "Language Development" in the 1960 edition of the *Encyclopaedia of Educational Research*.

words per clause. "Subordinate clause index" is the number of clauses per T-unit. "Average T-unit length" is the number of words per T-unit. "Main clause coordination index" is the number of T-units per sentence. "Average sentence length" is the number of words per sentence.

These five measures are very useful analytically and are all related arithmetically. The number of words per clause times the number of clauses per T-unit equals the number of words per T-unit. That times the number of T-units per sentence gives the number of words per sentence. The first index times the second equals the third. The third times the fourth equals the fifth. Clause length times subordinate clause index equals T-unit length. That figure times main clause coordination index equals sentence length.[7]

Finally, as a review, let me mention again the tendencies that have been known for thirty years concerning the development of language structure. First, as students mature they tend to have more to say about any subject. Second, as students mature their sentences tend to get longer. Third, as students mature they tend to write more subordinate clauses per main clause.

In the last few years a few more statements about syntactic development have been added. First, as students mature they tend to produce longer clauses. From kindergarten to at least the seventh grade, and probably beyond that time, this appears to be true of speech. And from the beginning to at least the twelfth grade this appears to be true of writing. Clause length is a better index of language maturity than sentence length. You will recall that clauses can be lengthened by a process that is here described as reduction and consolidation. That same process is described by generative-transformational grammarians as embedding transformations. Second, a convenient unit, intervening in size between the clause and the compound sentence is the "minimal terminable unit," defined as one main clause plus whatever subordinate clauses are attached to it or embedded within. This too is a better index of language maturity than sentence length. You will recall that "T-units" can be lengthened either by lengthening clauses or by increasing the number of subordinate clauses per T-unit. In the writing of average students throughout the public school grades, the one factor is about as influential as the other in effecting longer T-units. But the equality of influence stops there. The average twelfth grader has approached the ceiling in number of subordinate clauses. To advance beyond the level of the average twelfth grader, the writer must learn to reduce and consolidate clauses much more often. Superior twelfth graders do not write more subordinate clauses than average twelfth graders. Instead they write much longer clauses, just as superior adults do.

But making one clause out of two is child's play. Long before the average child gets to the fourth grade he can consolidate two, though he does not do so very often. Some average fourth graders consolidate three into one. Some average eighth graders consolidate four into one. Some average twelfth graders consolidate five into one. Superior twelfth graders consolidate six and seven. Superior adults consolidate more than that.[8] For a teacher to stretch a young-

[7] Kellogg W. Hunt, "A Synopsis of Clause-to-Sentence Factors," *English Journal*, 54 (April 1965) 300-309. Also Kellogg W. Hunt, *Grammatical Structures Written at Three Grade Levels*, Research Report #3, NCTE, Champaign, 1965.

[8] Kellogg W. Hunt, *Sentence Structures Used by Superior Students in Grades Four and Twelve, and by Superior Adults*. USOE Research Project No. 5-0313. Available from ERIC Document Reproducing Service, Bell and Howell, 1700 Shaw Boulevard, Cleveland, Ohio 44112.

ster, to push against the limits of his present accomplishment, two is nowhere near enough.

So far I have talked only about building up little sentences into bigger ones. Before I close I want to mention the other side of the coin, breaking big sentences down into little ones, as the mature reader or listener does with such lightning speed.

In recent months we have compared the syntactic traits that make a sentence hard to write and the syntactic traits that make it hard to read. We have compared the sentences written by children with those read by them but written by adults. For our reading samples we have used the *McCall-Crabbs Reading Lessons,* since the readability of each passage therein is supposedly already established. The passages cover roughly grades four to nine.

We find that sentences more difficult to read do not have more T-units per sentence. The number is about the same whether the sentences are easy or difficult. Listen to this sameness for grades 4 to 9: 1.13, 1.12, 1.13, 1.10, 1.13, 1.10.

You will remember that as children mature they tend to write subordinate clauses more often. But as sentences written by adults get easier or harder to read there seems to be no change. Listen to this sameness for passages that are 80 percent comprehensible to children in grades 4 through 9: 1.4, 1.4, 1.4, 1.4, 1.4, 1.4 (It sounds as if the record player were stuck in the same groove.)

But the clause length of passages, for 80 percent comprehension, increases steadily for grades 4 through 9 just as it increases as children write. Here are our figures so far: 8.45, 9.13, 9.59, 10.19, 11.01, 10.83.

In other words, the difficulty in reading sentences usually lies down inside the clause. Longer clauses tend to be more difficult. On the basis of previous research on what constitutes these clauses, it seems clear what it is that makes longer clauses harder to read and harder to write. On the whole, longer clauses have a larger number of sentences or clauses reduced and consolidated into one. It is by that process that little clauses grow into big ones.

Little by little the evidence piles up that the reduction and consolidation of many clauses into one is intimately related to syntactic growth both in writing and reading. If writers must build up clauses, then readers must break them down. A whole new range of applications is opened up for approaching reading difficulty.

For many years we have known that longer sentences tend to be harder to read. But in the last few months we have learned more about why that is true. It is not because longer sentences have more T-units coordinated into them, for they do not. It is not because they have more subordinate clauses attached to main clauses, for they do not. Instead it is because the clauses are longer. And the clauses are longer, it can be inferred, only because more have been consolidated into a single one.

All this has implications for the teaching of reading in the early grades. Teachers need to be trained in clause-consolidation so that children can be taught what otherwise they must discover unaided. They will discover it. That we know. But at present they must do so unaided.

Here is another place where the results of research should crawl out of the learned journals and into the classroom.