Both sentence length and sentence complexity of various kinds of newspaper prose were examined in a study to determine (1) ways newspaper prose differs from the prose of books native English speaking students chose to read and enjoy reading and (2) whether there are features of newspaper prose other than vocabulary and a lack of international news keeping advanced second language students from reading the news. A random sample of 100 sentences each was taken from two major Philadelphia newspapers, along with a sample of 100 sentences each from two of the major international wire services. In addition, a random sample of 25 sentences each was taken from a sample of 8 books chosen by community college students. These sentences were then compared for sentence length and syntactic constructions, first to identify the differences between them, and second to formulate reading strategies to cope with the special problems presented by the newspaper prose. The analysis revealed that newspaper syntax differed from the book prose in the number rather than the kind of construction, with each construction appearing more frequently in the newspaper prose. Taken as a group, however, the constructions resulted in systematically different sentences. (The paper includes a listing of reading strategies for reading the newspaper. The appendices list the sample methodology and the book prose samples.) (HOD)
Coping with Newspaper Syntax: Reading Strategies
for L1 and L2 Learners

Barbara Hoekje
Coping with Newspaper Syntax: Reading Strategies for L1 and L2 Learners

Reading the newspaper is for many an ideal leisure time activity: "I'm going to go home, read the newspaper and relax." Newspaper writing is generally considered to be popular writing, easily accessible to the general public, and standing in contrast to more difficult reading such as academic textbooks.

Yet for some populations, reading the newspaper is more work than relaxation. In a basic reading skills course at the Community College of Philadelphia, students were polled on the extent to which they read the newspaper, and which sections they enjoyed the most. The course the students were taking is designed for students who have been admitted to the college but are required to take a course to improve their basic reading skills before taking credit college courses. The typical student is an L1 learner graduating from the Philadelphia public school system; there are also a few L2 students with higher-level English skills than those who are placed in the regular ESL component of the English Department at the college.

The students reported that when they read the newspaper at all, that certain feature columns and the sports page were by far the overall favorite sections. Front page news and editorials were nobody's favorite sections, and by self-report read so infrequently that they could hardly be counted as a significant source of information for these students at all. This was not an atypical class by any means: students in this basic reading skills class in the past have responded similarly. The students said that they "just didn't like" to read the news stories in the newspaper; they were interested
enough in the events of the world but they preferred getting
the news on t.v. When reading the news articles in class the
students stumbled at every point: unfamiliar vocabulary, dense
syntactic constructions, insufficient background knowledge.
Reading the news stories was a difficult enough process that it was
clear why these students just "didn't like" to pick them up and
read them on their own. Yet these were students with a great deal
of interest in the events of the world who were also reading outside
material during the course and enjoying it. Each student read
several books that he or she selected and reported orally on.
Comprehension of these books was good and enjoyment was generally
high.

Other populations such as L2 learners also typically have
problems with reading the newspaper long after they are familiar
with other types of prose. L2 learners in the most advanced reading
class (Level 700) at the English Program for Foreign Students at
the University of Pennsylvania were also polled on the extent to
which they read the newspaper. These students in general also
avoided the newspaper, they said because it didn't carry much
international news. When asked what problems if any they had reading
the newspaper, they unanimously stated that vocabulary was their
major problem. None of these students was doing any reading in
English at the time other than that required by the course;
however, there was a wide variety of other prose reading in the
course that the students completed regularly and without much
difficulty.

The responses of these two groups of students, both reading
some kinds of prose with enjoyment, both avoiding the front page
news stories, raised the question of readability: exactly how readable is the newspaper? In what ways did newspaper prose differ from the prose of the books the L1 students chose to read and enjoyed reading? Were there features of newspaper prose other than vocabulary and a lack of international news that could be keeping the advanced L2 students from reading the news?

**Readability**

Previous measurements of readability have relied primarily on word length of sentences and difficulty of vocabulary: that is, the percentage of unfamiliar words as determined by a word list (Harris and Jacobson, 1973). Fry's (1968) simplified reading formula uses only word length of the sentence and the number of syllables/word to predict readability: the results from his formula correlate highly with other results from the SRA, Dale-Chall and Flesch studies.

Sentence complexity has generally been treated in readability studies as a function of sentence length (Fry, 1968:576). Research in first language acquisition also has used mean length of utterance as an indicator of sentence complexity (Newport, Gleitman and Gleitman, 1977).

In this study both sentence length and sentence complexity were examined in an attempt to identify the number and kind of syntactic constructions appearing in various kinds of prose.
The study

A random sample of 100 sentences each was taken from two major Philadelphia newspapers, the Philadelphia Inquirer and the Philadelphia Daily News, along with a sample of 100 sentences each from two of the major international wire services, United Press International (UPI) and Associated Press (AP). The UPI and AP articles were chosen as a sample because they are carried in newspapers throughout the nation. This total of 400 sentences from the newspapers will be referred to herein as the newspaper sample. A random sample of 25 sentences each was taken from a sample of eight books chosen to read by the students at the Community College of Philadelphia, with a total sample size of 200 sentences. This sample will be referred to herein as the book prose sample. The book prose sample is small, and useful here primarily for comparison purposes; this study does not purport to give a comprehensive description of the book sample prose.

The two sets of sentences, those from the newspaper and those from students' self-selected books, were compared for sentence length and syntactic constructions, first in order to identify the differences between them, and second to formulate reading strategies to cope with the special problems presented by the newspaper prose.

Methodology

The appendix contains the complete information about the sampling and coding methodology used in this study. The major considerations in the sampling were first, to get sentences from the beginning, middle, and end of the articles and books to avoid any stylistic
bias, and second, to choose front page news stories rather than sports stories or features, since it is precisely the syntax from news stories that is under consideration here.

Consistency

The Education Department of the publishing company that owns both the Philadelphia Inquirer and the Daily News could give no official estimate of the readability level of either paper, although their impression was that the Daily News was written at an easier level. In fact, the results for the four newspapers sampled in this study are highly consistent with each other in the variables measured here.

The prose sample, taken from eight books randomly chosen from a much larger sample of books (see Appendix for complete list) shows less consistency. One book in particular, Paul Robeson: Biography of a Proud Man has characteristics closer to the newspaper prose style than to the book prose in this sample. This book is a biography and is written in a journalistic style very comparable to the newspaper style. In considering the overall averages of the book prose sample, the uncharacteristic behavior of Paul Robeson should be taken into account in every result.
A comparison of the two styles of writing shows that newspaper syntax differs from the book prose in the number rather than the kind of constructions, with each construction reported on here appearing more frequently in the newspaper prose. Taken as a group, however, the constructions result in systematically different sentences.

Sentence Length

An examination of sentence length shows a significant difference between the two styles (see Figure 1). The average length of newspaper sentences is 24 words per sentence, while the average length of the book prose is slightly more than half that: 14.4 words per sentence. 115 out of the 400 sentences had thirty or more words; the longest sentence was fifty-seven words long. This length is partially the result of a number of compound sentences,
that is, several independent clauses connected together by
and, but, or, for or nor, where the reader finishes one independent
thought before moving on to the next. Out of a sample of 400
sentences, 34, or 8% in the newspaper sample were compound. The
rest of the sentences; 92% of the sample, consisted of sentences
with one independent clause which was usually surrounded or em-
bedded with one or more dependent clauses. In the book prose
sample, the number of compound sentences is slightly higher:
23 out of the 200 sentences sampled, or eleven percent.

As discussed above, earlier studies measuring readability have
relied primarily on sentence length and vocabulary as an index;
this study confirms that sentence length is integrally connected
with sentence complexity.

Subject Noun Phrase

The subject noun phrase (NP) of each independent clause was
analyzed for length (See Figure 2). The subject NP's in newspaper
prose were twice as long, with an average of four words per subject,
as those in book prose. Subject NP's in the newspaper sample ranged
from one to 23 words in length: e.g., "Former Green Beret Officer
James "Bo" Gritty, whose Hollywood connections helped finance a
mission into the Laotian jungle in search of American POW's, sur-
rrendered today to Thai police." (UPI: Daily News 2/28/83 p. 20)

The length of the subject noun phrase comes from both pre-head
noun modifiers such as "Liquor Control Board Chairman Daniel Pennuk
says...", and post-head noun modifiers in the form of participial
phrases, relative clauses, prepositional phrases or appositive phrases, such as "The Reagan Administration, taking its tough stand on crime to the Supreme Court, says..." (UPI 2/14/83 p. 6). Readers must learn to expect a noun phrase greater in length than that occurring in other kinds of prose, with modifiers occurring both pre- and post-head noun. It will very frequently happen, therefore, that the subject noun will not be immediately followed by its main verb.

Very large NP's occur in all positions in the sentence, not just in subject position. More than features of the verb or other syntactic constructions, it is this characteristic of newspaper prose, a large "packed" noun phrase, which has also been noted as the major distinguishing characteristic of more academic formal prose (Hunt, 1970).
Number of Propositions/ Sentence

The definition of a proposition for the purposes of this paper is the linguistic encoding of an event into its logical subject and verb. Thus "Mary saw an accident" is one proposition; "Mary called the police" is one proposition; "Mary saw an accident and called the police" is two propositions, and "Mary, who lives at 4329 Walnut St., saw an accident and called the police" is three propositions.

The number of propositions per sentence was investigated in this study (see Figure 3). The average number of propositions per sentence in newspaper prose is about two and one-third, which is seven propositions in every three sentences, compared to about one and three-fourths per sentence in the book prose.

In addition to the propositions encoded by a grammatical subject and verb, there are propositions encoded in the noun phrase such as "The President's abhorrence of tax increases is well known." This
noun phrase, "The President's abhorrence of tax increases" contains the proposition "The President abhors tax increases". Another example is the following: "First time claims for unemployment benefits by jobless Americans totaled 491,000 in the week that ended January 15, the lowest filing level since September, 1981, the Labor Department reported yesterday." The subject NP is "First time claims for unemployment benefits by jobless Americans"; this contains the proposition "jobless Americans claimed unemployment benefits for the first time".

When the number of propositions per sentence was being counted, this type of proposition was not included; however, NP's containing propositions of this form are relatively frequent in newspaper prose, occurring approximately in 10% of the sentences.

Passive

Verbs appeared in the passive much more frequently in the newspaper sample than in the book sample. In fact, were Paul Robeson excluded from the book sample, the passive would appear only twice in 298 verbs, or seven-tenths of one percent (.7) in the book sample.

This is a rather unexpected result since the passive (with be or get) is not an uncommon form in speech or informal writing. In this sample of both newspaper and book prose, however, it occurs only with be (never with get), and six times as frequently in the newspaper sample as in the book sample, even including Paul Robeson, which is a book highly uncharacteristic of the book sample.

This has implications particularly for L2 readers. Since an overall average of one out of every eight inflected verbs they will be encountering in the newspaper will be in the passive, they need
to be prepared for its frequent occurrence, alone or sharing the subject with an active verb in conjoined verb constructions such as "Borders was convicted of bribery and conspiracy by an Atlanta court last summer and is appealing a five-year prison term." (UPI: Philadelphia Inquirer 2/2/83).

Embedded Sentences

The frequency of embedded sentences in the newspaper prose sample was twice that of embedded sentences in the book prose sample (see Figure 4). The sentence embedding in the newspaper came predominantly after verbs of saying such as say, tell, report, contend, for example, "Vice-President Bush said yesterday that there would be no deferral of the NATO plan to deploy new missiles in Europe". Embedding also frequently occurs after a noun made from a verb of saying, such as statement, report, allegation, and so forth, for example, "Church of Scientology officials yesterday told reporters allegations that the Church founder is dead were an "extortion attempt." (AP: Inquirer, 2/11/83 A1).

![Figure 4: Number of Embedded Sentences Per Matrix Sentence](image-url)
In newspaper prose, there is the additional complication that the subject of the matrix sentence is often heavily modified, providing a great deal of new information in the proposition of the matrix sentence at the same time as the reader is asked to cope with a minimum of two different propositions, that of the matrix sentence and that of the embedded sentence. Embedding also frequently occurs without the embedding conjunction that; thus processing of the information as an embedded clause is made more difficult (cf. the example above, "Church of Scientology officials yesterday...told reporters (that) allegations that the Church founder is dead...").

Relative Clauses

Relative clauses occur about once in every four independent clauses in newspaper prose (see Figure 5). They function non- restrictively, giving additional information about the head noun, as in "This could turn into a free-for-all, said Randol, who analyzes U.S. oil companies for the investment firm of First Boston Corporation." (AP: Inquirer 2/2/83)
They also function restrictively, giving information necessary to identify the head noun, as in "The Syrians have been in Lebanon for 7 years, ostensibly to police the armistice that ended the 1975-76 Muslim-Christian civil war." (AP Inquirer 1/28/83 B-C)

The frequency of non-restrictive relative clauses in this data supports earlier observations about the substantial role played by non-restrictive relative clauses in newspaper prose (Dayton, 1980). It is reasonable to assume that this would be so: relative clauses are used as a way to add another proposition about the head noun, not necessarily for purposes of identification.

Dependent Clauses and Participial Phrases

Dependent clauses occur approximately once every five independent clauses (see Figure 6). About three-fourths of the time they appear after the verb phrase. Twenty-eight percent of the time that they do appear, they appear before the subject NP. When this happens, the number of words occurring before the subject NP can be large. Up to 28 in this sample (cf. Inquirer 1/28/83: "Angered by Taxes"). When dependent clauses do appear in pre-subject position, readers need to be prepared to receive a substantial amount of information before the independent clause.

![Figure 6: Number of Dependent Clauses Per Matrix Sentence](image)
Participial phrases occur more frequently than dependent clauses, about once in every four independent clauses (see Figure 7). They occur in both active form, as in "The Soviet Union, ignoring complaints that it has been interfering in next month's West German elections..." (UPI Inquirer 2/21/83 A), and in passive form, as in "a swelling crime rate fed by immigrant youths..." (UPI Inquirer 2/27/83 4C). Participial phrases occur both pre- and post subject position of the independent clause.

Traditional grammars have held that dependent clauses contain information less important than the independent clause, that a dependent or subordinate clause was simply that, subordinate to the information of the independent or main clause. The use of dependent clauses and participial phrases in newspaper prose appears to function differently, generally carrying information of equal importance to that carried by the independent clause.
In the English Program for Foreign Students at the University of Pennsylvania, students in the most advanced reading class were asked to identify the main subject and its main verb in a front page news article about the Pope in the Inquirer. The first two sentences in the article followed the same pattern: a subject noun phrase followed by a participial phrase and then the main verb. For both sentences, students identified either the participle, or the main verb, or both verbs as being the main verb in the sentence. The participle was identified as the main verb of the sentence as often as the main verb was.

One explanation for the students' choosing the participle as often as the main verb might be that readers look for the first verb following the subject NP and assume that that is the main verb of the subject. A second explanation is that the information introduced by the participial phrases was important and of as much relevance to the story as the information in the main verb.

Dependent clauses, relative clauses, and participial phrases surround the independent clause, not as constructions carrying necessarily subordinate ideas, but as syntactic devices to add more information and more propositions as compactly as possible.

The information carried in a dependent clause and a relative clause can be examined both in terms of its identity as a single proposition and as a proposition with a relationship to the rest of the sentence. Readers, especially L2 readers, need to be made familiar with the forms of participial phrases, especially passive participial phrases, and their relationship with a head noun.
Appositives

The rarest of the constructions that were examined is the appositive construction—a noun phrase that is set in juxtaposition to another noun phrase and set off by commas. Only fifteen appositives appeared in the 223 sentence sample of book prose (6.7%) (see Figure 8), while 61 appositives appeared in the 434 independent clauses in newspaper prose (14%), meaning that an appositive appeared more than twice as frequently in newspaper prose as in book prose, in about one out of every seven independent clauses.

This is one construction in particular that readers need to be familiarized with—both in terms of its form and its function—since they are not likely to meet it as frequently in other kinds of prose.

Prepositional Phrases

Prepositional phrases are the most common of the constructions examined here, appearing with a frequency of almost two and one-fourth per independent clause in newspaper prose and about half that in book prose (see Figure 9). Prepositional phrases function
Prepositions occur in every place in the sentence: sentence-initial as in "In another development concerning the EPA, fired agency official Rita Lavelle..." (UPI: Inquirer 2/27/83 Al); sentence-medial, often between subject head noun and verb, as in "A family marooned for nearly a month on a sand bar in the South Pacific returned home yesterday." (UPI: Daily News 1.14/83); or post-verb or sentence-final as in "About 10:15 p.m., police found the pickup truck abandoned nearly two miles from the shooting scene in a dense area of Cobbs Creek Park near West Chester Pike and State Road in Upper Darby Township."
Prepositional phrases as a construction are not likely to cause major difficulties for L1 readers, except 1) when they occur sentence-initially and provide substantial information before the subject and verb of the independent clause, and 2) in adding length and information to the noun phrase and making it more complicated to process. For L2 readers, prepositional phrases, especially those in phrasal combination with the verb or with relatively unfamiliar prepositions may cause reading difficulty; however, the vast majority of prepositional phrases in both samples were those beginning with the familiar prepositions in, of, by, at and on.
Conclusion

Understanding the syntax of individual sentences is, of course, only one part of the reader's task of understanding newspaper prose. The discourse, properties of newspaper prose, the structure of old and new information, shared knowledge and the marking of theme are all relevant to the task of reading competently. Familiarity with vocabulary or strategies to acquire new vocabulary are also necessary. Finally, the reader's extralinguistic knowledge of the world, which is relevant to the writer's assumptions about shared knowledge, is also a factor in a reader's ability to read the newspaper competently.

Nevertheless, news story prose clearly has special syntactic characteristics which distinguish it from other prose and which may present problems for even those L1 and L2 students who can read and enjoy other prose. Reading teachers must prepare students for front-page newspaper reading as carefully as for any other specialized reading tasks. In the beginning this will involve painstaking dissection of sentences into their component parts, and discussion of the relationship of the parts to each other. It will also involve presentation and practice with some of the less frequent syntactic constructions. Although there is no one set formula for the syntax of newspaper sentences, there are patterns that will become familiar to the students over time. The following reading strategies in particular have been useful in increasing the interest in reading and self-confidence of both L1 and L2 students:
Intensive Reading Strategies for Reading the Newspaper

1. Have students practice processing large noun phrase units with both pre- and post-head noun modifiers.

2. Have students practice processing noun phrases with embedded propositions such as "The President's abhorrence of tax increases is well known."

3. Discuss ways of coping with the length of the sentences; identify resting places when they are syntactically available as in compound sentences; use clause boundaries as resting places even when the clauses are syntactically dependent.

4. Prepare students to expect several propositions in each sentence encoded in constructions which may be syntactically subordinate but not necessarily informationally subordinate to the independent clause; practice identifying the propositions and their relationship with each other.

5. Be sure that students, especially L2 readers, are familiar with the passive form and can identify its subject, especially when it is linked with an active verb.

6. Be sure students are familiar with the form of both active and passive participial phrases and can identify their head noun.

7. Familiarize students with the form and function of the appositive construction.

8. Have students practice processing embedded constructions after both verbs and nouns of saying (She refuted the allegation that...)

9. Discuss the role of prepositional phrases in identifying place and time and providing information at any place in the sentence.

10. Prepare students to expect a substantial quantity of information before the subject noun phrase or between the subject noun phrase and its verb.


APPENDIX A: SAMPLE METHODOLOGY

The articles in the newspaper prose sample were chosen as follows:

I. The Philadelphia Inquirer

All the front page articles that were written by the Inquirer Wire Service or Staff (i.e., not the UPI or AP-written articles) were selected from randomly selected issues during the months of January, February, and March, 1983, up to the required number of thirty articles.

II. The Philadelphia Daily News (magazine format)

The first five articles on the odd pages (odd pages only to avoid logistic problems when clipping the articles) written by the Daily News wire services or staff were chosen from randomly selected issues during the months of January, February and March, 1983, up to the required number of thirty articles.

III. AP and UPI

Any articles written by the AP or UPI which appeared in the Inquirer or the Daily News in the issues used above were selected, with the exception of sports articles. Collection of articles stopped after thirty articles each were collected.

SELECTION OF SENTENCES

I. Newspaper sample (all sources)

Five sentences were chosen from each article, starting with the first sentence and then proceeding with every fifth sentence; thus the first, sixth, eleventh, sixteenth and twenty-first sentences were taken. If the article was not long enough to get all five sentences, the same method was used and fewer sentences were collected.

II. Book Prose Sample

The first five sentences on each of the following pages of each book were selected, for a total of twenty-five sentences from each book: the tenth page from the beginning, the tenth page from the end, three pages equidistant from each other between the first tenth and the last tenth page.
APPENDIX B: THE BOOK PROSE SAMPLE

Sentences were randomly selected from the following eight books:

7. Steinbeck, John. Of Mice and Men

The above books were randomly chosen from the following 17 books selected by students in a basic reading skills class at Community College of Philadelphia, fall, 1982.

1. Love at Forty
2. The Empire Strikes Back
3. Captain Flamingo
4. Passing
5. Of Mice and Men
6. P.S. I Love You
7. The Amityville Horror II
8. The Golden Heal
9. Paul Robeson: The Biography of a Proud Man
10. Promises in the Dark
11. They Were Expendable
12. If Beale Street Could Talk
13. Sybil
14. The Dead Zone
15. Prison to Praise
16. Other Side of Midnight
17. The Pearl