

ON THE DIFFERENCES BETWEEN SPOKEN AND WRITTEN LANGUAGE*

F. NIYI AKINNASO**
University of California, Berkeley

Drawing on research studies in (socio)linguistics, discourse analysis, and literacy, this paper provides a synthesis of findings about lexical and syntactico-semantic differences between spoken and written language, focusing on empirical research on the English language since the 1920s. The major theoretical and methodological approaches used in comparative studies of spoken and written language are outlined and their advantages and shortcomings are critically examined. The question of how speech and writing relate to prototypical forms of language and meaning is reexamined in the light of the findings. To avoid basing generalizations about differences between speech and writing uniquely on English, suggestions for future comparative research are offered and discussed.

1. INTRODUCTION

Since the turn of the century, the nature of the relationship between spoken and written language has been a subject of considerable interest in linguistics, anthropology, education, and psychology. While anthropologists are primarily interested in the implications of writing for cultural evolution and the growth of social systems (cf. the equation, "civilization = writing" [Goody, 1980, p. 120]), educators and psychologists have focused on the cognitive factors affecting its acquisition. Linguists are interested in several aspects of the relationship between speech and writing according to their theoretical perspectives. For example, psycholinguists are almost neatly divided into two distinct groups, one looking at the processes of oral language acquisition (e.g., Brown and Bellugi, 1964), and the other at the neuropsychology of reading and writing processes (e.g., Guthrie, 1976). Socio- and applied linguists have looked at more practical aspects of speech, reading, and writing activities, emphasizing the contexts of those activities and their consequences for the learners or participants (Cook-Gumperz and Gumperz, 1981; Szwed, 1981). Linguists and language teachers interested primarily in the structure of language have focused on the lexical and grammatical differences between spoken and written language (Drieman, 1962; O'Donnell, 1974). Recently, discourse analysts, students of literacy, and language teachers have shown a keen interest in both the differences and similarities between spoken and written language (Golub, 1969; Poole and Field, 1976; Ochs, 1979; Chafe, 1979, 1982; Goody, 1977, 1980; Tannen, 1980;

* *I am grateful to John Gumperz and Wallace Chafe for comments on an earlier version of this paper and to the editorial readers for this journal for helpful suggestions.*

** *On leave from the Department of Linguistics, University of Ife, Nigeria.*

Stubbs, 1980; Lakoff, 1981; Akinnaso, 1981a, 1981c).

There are at least three reasons for the current upsurge of interest in the relationship between spoken and written language. One is the contemporary significance of studies of the cognitive, linguistic, and sociocultural consequences of literacy (see Akinnaso, 1981b for a review). This has meant closer attention to the differences between oral and literate traditions, focusing on the possible differences in communicative strategies and in modes of thought between the two traditions. Since it is one of the major variables in the dichotomy, writing has had to be examined more critically in relation to speech. Drawing on a wide range of historical, philological, and ethnographic sources, Goody's work (especially, Goody, 1977) represents the most provocative treatment of the communicative and sociocognitive import of writing so far.

A second factor is the current interest in discourse analysis and "communicative competence" which extends the domain of linguistic analysis beyond both the sentence and mere "linguistic competence" to include performance data drawn from both spoken and written language samples. Earlier analyses of linguistic data had focused almost exclusively on either spoken or written language. Historically, before the advent of writing (and especially of printing), whatever "linguistic" analysis existed was carried out under the rubric of Rhetoric or the transmission of ritual texts (as in India), and was based exclusively on oral data. Printing and widespread uses of writing later came to arouse interest in written, especially literary, data, shifting attention from oral language. However, attitudes changed again by the end of the nineteenth century (and especially during the first two decades of the twentieth) when in Germany, the Grimm brothers were recording the speech of illiterate peasants; in Britain, Henry Sweet and Daniel Jones were developing phonetic studies; and in the United States, American structuralists were turning their attention to the largely unwritten American Indian languages, dismissing writing as "not language, but merely a way of recording language by means of visible marks" (Bloomfield, 1933, p. 21). Although, after a large body of oral linguistic data had been recorded and described by the structuralists, American transformationalists came to dismiss spoken language as mere performance, too random and unsystematic for serious study (Chomsky, 1965). Sociolinguists and discourse analysts have, in the past two decades, demonstrated the need and validity of drawing data from both spoken and written language. However, in accepting both spoken and written language data, researchers have had to evaluate the relationship between the two modes of communication in order to determine which sorts of data are acceptable for study. Recent work by Keenan and Bennett (1977), Ochs (1979), Chafe (1982), Stubbs (1980), Tannen (1980, 1981), Gumperz, Kaltman, and O'Connor (1982), and Akinnaso (1981a) demonstrates this awareness, in various ways.

Another reason for the upsurge of interest in the relationship between spoken and written language stems from current concerns with problems of transition to literacy and the ongoing debate concerning the "linguistic discontinuity hypothesis," whether conceived in terms of Bernstein's code elaboration model (Bernstein, 1961, 1964, 1971), in terms of variation in dialect or context (Labov, 1972a, 1972b), or in terms of "mismatch" in communicative etiquette (Philips, 1972, 1975; Simons, 1979). The hypothesis is based on the dual assumption (1) that there is linguistic incompatibility

between home and school, especially on the part of minority and lower-class children who are believed to speak an "oral" language as opposed to middle-class children who are believed to "speak a written language" (Greenfield, 1972), and (2) that this incompatibility is a major contributory factor to the educational under-achievement of minority or lower-class children. Within the past two decades, studies have been conducted by sociolinguists, anthropologists, educators, and psychologists to test the validity of these assumptions, and to account for the communicative problems involved in the transition to literacy (Lawton, 1963; Greenfield, 1972; Poole and Field, 1976; Olson, 1977a, 1977b; Goody, 1980; Cook-Gumperz and Gumperz, 1981; Collins and Michaels, 1980; Wells, 1981; Gumperz, Kaltman, and O'Connor, 1982).

However, in spite of the vast amount of data that have been accumulated on the relationship between spoken and written language, no agreement has been reached either on the exact nature or the extent of the differences (and similarities) between the two. While it is true that many of the studies have produced evidence of superficial differences between speech and writing, several findings on the proportionate distribution of certain lexical categories and syntactic features have been conflicting. Furthermore, a number of researchers have stopped at mere taxonomic or typological differentiation between spoken and written language without explaining why the observed differences occur, or how knowledge of such differences can be used to guide learners in the transition from speech to writing.

The purpose of the present paper is to provide a synthesis of major findings about lexical and syntactico-semantic differences between spoken and written language, focusing on empirical research on the English language. Although some attention is given to the historical development of research on the subject, the emphasis of the review is on the theoretical and methodological bases of the studies and their pedagogical implications. Section 2 of the paper focuses on lexical differences between spoken and written language, and Section 3 is devoted to studies of syntactico-semantic differences between the two modalities. In Section 4, some of the major explanations given for the surface differences between spoken and written language are reviewed, and Section 5 is devoted to a discussion of major shortcomings of traditional explanations and a review of a new approach which examines the relationship between spoken and written language within the broader context of the acquisition of literacy. Section 6 offers a summary and discussion of major theoretical views on how speech and writing relate to prototypical forms of language and meaning, and concludes with suggestions for future research on comparative studies of spoken and written language.

2. VOCABULARY CHOICE AND LEXICAL STRUCTURE

Since the early 1920s, several studies have used word frequency counts as a primary method for distinguishing between spoken and written language (e.g., Dewey, 1923; Horn, 1926; Thorndike, 1932; Voelker, 1942; Fossum, 1944; Drieman, 1962; Blankenship, 1962; DeVito, 1964, 1965, 1967; Gibson, Gruner, Kibler, and Kelly, 1966; Gruner, Kibler, and Gibson, 1967; Chafe, 1982). Although most researchers agree that spoken

and written English differ in word choice and lexical structure, there are major differences in (1) data base; (2) method of data gathering and analysis; and (3) the specific purpose of word count.

The earliest studies began by looking at spoken and written language separately without systematically correlating the findings. For example, Horn (1926) attempted a census of 10,000 words "most commonly used in writing," while Voelker (1942) counted what were considered "the one thousand most frequent spoken words." Another pair of studies that looked at spoken and written language separately is offered by Fairbanks (1944) and Bachman-Mann (1944). Although these workers used spoken and written language data respectively, their concern was more about the differences in linguistic behavior (in speech and in writing) between schizophrenic patients and "normal" subjects rather than the differences between spoken and written language per se. One interesting technique used by Fairbanks and Bachman-Mann, and which has been found useful in quantitative studies of the relationship between spoken and written language, is the computation of Type-Token-Ratios (TTR) -- the ratio of the number of different words (types) to the total number of words (tokens) in a sample of language data. For example, if in speaking or writing 100 words, an individual uses 72 different words, his TTR would be 0.72. TTR can be computed for any word class or grammatical category and the results can be easily used for comparative purposes.

Building partly on work by Fairbanks and Bachman-Mann, Drieman's (1962) much cited study provides perhaps the best reference point for the study of lexical differences between spoken and written language, partly because of its methodology, and partly because of its many interesting findings. Drieman's research rested on four fundamental assumptions: (1) that comparative data on spoken and written language should be based on an identical topic for the spoken and the written protocols; (2) that data should be obtained from the same subjects; (3) that the circumstances under which data are obtained should be identical for all subjects; and (4) that, as different segments of a discourse contain varying amounts of detail, "only the *entire* oral and the *entire* written communication are comparable" (Drieman, 1962, p. 39; original emphasis). Pretending that he was interested in psychological rather than linguistic processes, Drieman showed two pictures to a group of eight graduate students in psychology and elicited their spoken and written responses. The instruction was that the subjects evaluate the pictures and describe their effects on them. This was later followed by an orally administered questionnaire, requiring the subjects to evaluate their spoken and written samples. In many ways, the results of this qualitative study corroborate those of the quantitative study, attesting subjects' awareness of possible differences between their spoken and written language.

In the quantitative study, the following features were found to be diagnostic of written language: (1) longer words (i.e., more polysyllabic words); (2) more attributive adjectives; (3) a more varied vocabulary; and (4) shorter text. Drieman was, of course, aware of some of the limitations of this study. Variables which may have affected the findings include the number and educational level of the subjects, the nature of the communicative task specific to the stimulus situation, and the brevity of texts (see Drieman's appendix). Despite these limitations, however, Drieman's findings corroborate

others; indeed, one interesting aspect of the study is Drieman's comparison of American, French, and Dutch studies using similar procedures. The comparison shows that the results were similar despite considerable variations in choice of subjects, experimental setup, and specific rules of differentiation between spoken and written samples. Two major reasons for the similarity of findings are (1) similarity of educational background for all subjects, and (2) employment of the same measurement technique, computation of TTR for specific word classes.

Gibson *et al.* (1966) improved on Drieman's method with two other measures: Flesch Reading Ease (average sentence length and average number of syllables per 100 words) and Flesch Human Interest (based on the number of personal words and sentences directly addressed to the reader). Again, subjects were university students; half of them were required to speak and later write on a topic chosen from a pool of alternatives, while the procedure was reversed for the other half. As each subject chose a different topic, the data accumulated turned out to be series of individual cases, each subject being involved, more or less, in a different communicative task. Furthermore, unlike Drieman, Gibson *et al.* limited analysis to the first 500 words in each subject's spoken and written samples. This notwithstanding, they corroborated Drieman's findings: "The spoken style tends to be characterized by fewer words, words with fewer syllables, shorter sentences, and more personal words than written style" (p. 450).

In a series of similar studies, DeVito (1964, 1965, 1966a, 1966b, 1967) found that spoken language is "significantly less abstract and contained more finite verbs and less nouns of abstraction than written language" (1967, p. 359). This means that writers tend to choose words that are "significantly more difficult" than they would use when speaking on the same topic. This finding is consistent with that of Green (1958) and Kaump (1940), the latter reporting that "more hard words are used in written than in oral language" (p. 241). In addition to a higher level of abstraction, DeVito also found that written language has "greater verbal density" and "greater density of ideas" (cf. Chafe, 1982). Interestingly, DeVito's findings about the greater proportion of abstract nouns in written language has important implications for earlier observations about the length of words in written language. The derivational history of most abstract nouns (e.g., lateness, friendship, agreement) entails the addition of one morpheme (e.g., -ness, -ship, -ment) to morphologically simpler forms (e.g., late, friend, agree), thus producing longer words.

One method which DeVito introduced into his analysis to test comprehensibility is the "cloze" procedure, involving the random or systematic deletion of words (usually every fifth word) from a text composed orally or in writing (see Taylor, 1953). Comprehensibility is measured by feeding the altered text to the subjects who produced the original (and sometimes to others) and assessing the degree of fit between the original version and the subjects' responses to the modified text. Such responses can be assessed for their absolute (same word), grammatical (same form class), or semantic (equivalent meaning) accuracy. Application of this procedure showed that spoken and written language did not differ significantly in ease of comprehension, although vocabulary was significantly more difficult for written than for spoken language. DeVito's findings may have, no doubt, been affected by the educational and linguistic background of his informants (they were

university professors of Speech and Theatre) and the formality of the topic on which the spoken and written samples were based (published academic papers written by the subjects).

Portnoy also applied the "cloze" procedure to test the comprehensibility of spoken and written language. Like DeVito, she found no significant difference. However, individual subjects in her study produced pairs of spoken and written passages that differed significantly in comprehensibility from one subject to another. Those whose productions were more comprehensible she called Speakers, and those who produced more difficult texts she called Writers. It is the latter who tended to use more difficult and longer words in both their spoken and written messages. Although it is not clear from Portnoy's study what factors may have contributed to the individual differences, her conclusions reinforce everyday observations about differences between the spoken and written productions of individuals: "Writers are not necessarily Speakers, nor Speakers Writers" (Goody, 1980, p. 125).

In their own comparative study of spoken and written vocabularies, Gruner *et al.* (1967) looked for both differences and similarities. They reported that although subjects used a more varied vocabulary in writing than in speech, the twenty-five most frequently used words for both spoken and written messages were similar (cf. Voelker, 1942; Fossum, 1944). The words included determiners, demonstratives, pronouns, and the conjunction *and*. The distribution of these words varied according to modality. For example, there were more personal pronouns in the spoken than in the written samples. This finding corroborates DeVito's earlier (1966a) observation that speech contains more self-reference words than writing. Also, because of its more egocentric style, speech contains more terms indicative of "consciousness of projection" than writing. Thus expressions like *I think, in my opinion, as far as I am concerned*, etc., which indicate that the observed is in part a function of the observer, tend to be eliminated in writing.

Gruner and his collaborators limited their study of lexical similarities between speech and writing largely to function words. As these have to be used regardless of variation in topic, participants, and modality, their findings are not particularly striking. Except for gross differences in word distribution, the more interesting problems of differences/similarities between speech and writing in the use of content words, especially basic level lexemes, have been largely ignored. Yet such study is important in deciding whether the "deployment of consciousness" or the factors influencing lexical choice are the same for speech and for writing (see Chafe [1980] and Downing [1980] for a discussion of these issues in regard to oral narrative production).

More recent researchers (e.g., O'Donnell, 1974; Kroll, 1977, Ochs, 1979, and Chafe, 1982) use word counts as the basis for making syntactic or discourse-related arguments. For example, Chafe (1982) counted the number of conjunctions occurring per thousand words in both the spoken and the written samples to show how their higher frequency of occurrence in the spoken protocols (66.0 vs. 15.8) leads to the greater "fragmentation" of spoken language. Similarly, in his attempt to account for the greater "integration" of written language, he counted, among others, participles used as nouns and verb nominalizations; he found that they were used more frequently in writing than in speech, as were several syntactic devices (to be discussed later).

While the above studies represent the major findings about the lexical differences between spoken and written language, they do not exhaust the possibilities. For example, everyday experience shows that spoken and written English differ in their usage of borrowed, classical, and technical vocabularies; but the nature of the differences has not been systematically studied. However, suggestive of the differences are a few studies that were not originally designed to compare spoken and written language. For example, through a series of experiments involving choice between synonym pairs, Levin, Long, and Schaffer (1981) have found that words whose etymology is Latin are favored over synonymous Anglo-Saxon words "under explicit instructions to be formal, or lacking such instructions, under conditions that are strongly coercive toward formal language" (p. 171). Thus Latinate words are more frequently used in writing than in speech, and more in formal than in colloquial speech. Similarly, technical vocabularies and word definitions are more frequently used in writing than in speech since they are devices commonly employed to enhance the autonomy and explicitness of written discourse. What all this implies is that the transition from spoken to written English involves, among others, a change in the pattern of lexical choice. Yet vocabulary drills, word recognition tests, and composition exercises given to beginning reading/writing students often show little or no awareness of such differences.

But change in modality is only one of several factors that may influence lexical choice. Even within the same modality, lexical choice may be affected by various factors, including (1) context and purpose of the speech event; (2) the nature of the communicative task appropriate to the speech event; (3) topic and associated register peculiarities; and (4) participants' background and level of linguistic knowledge. Therefore, in order to know which differences in lexical choice are due to change in modality (from speech to writing), all the variables listed above must be carefully controlled. The ability to control for some of them accounts, in part, for the success of Drieman's (1962) study.

Another problem with studies of lexical differences between spoken and written language, of which Drieman himself was a victim, is the limitation of the samples to the language behavior of a single segment of the population: academics (university professors and students) – the speech community to which the researchers belong. Largely ignored are elementary and high school populations where the greatest difficulty with written language is experienced. Although students and teachers of beginning writing as well as psychologists have investigated some of the problems that children encounter in the process of "translating" speech to writing and of decoding written language (Shaughnessy, 1977; Bartlett and Scribner, 1977; Goodman, 1972), the lexical differences between children's speech and their writing or their textbooks have not been systematically investigated.

A step in this direction is offered by Harrell (1957) and O'Donnell, Griffin and Norris (1967) – reviewed later – as well as recent work by Anderson, Armbruster, and Kantor (1980) which highlights a number of problems including the lexical and structural differences between adult and children's language and how such differences could be controlled in the writing of children's texts for classroom instructions. Such control tasks will remain difficult so long as models of children's written language are based on simplifications of adult models rather than on systematic comparisons of children's

spoken and written language data.

3. SYNTACTIC AND SEMANTIC STRUCTURE

From Woolbert's early observation that speech and writing differ in "syntax" and "sentence structure" (1922, p. 283) to recent empirical studies of syntactico-semantic differences between the two modes of communication (e.g., Keenan and Bennett, 1977; Ochs, 1979; Chafe, 1982), many interesting dichotomous distinctions have been drawn between spoken and written language with reference to the organization of ideas. Though there are notable contradictions in the findings, it is generally argued that, unlike speech, written language promotes the following:

1. Preferential usage of elaborate syntactic and semantic structures, especially complex nominal constructions (noun groups, noun phrases, nominalizations, relative clauses, etc.) and complex verb structures (Woolbert, 1922; Borchers, 1936; Drieman, 1962; DeVito 1964, 1966a, 1967; O'Donnell, *et al.* 1967; Huddleston, 1971; Poole and Field, 1976; Ochs, 1979; Chafe, 1982; Goody, 1980).
2. Preference for subordinate rather than coordinate constructions (Harrell, 1957; Blankenship, 1962; O'Donnell, 1974; Kroll, 1977; Ochs, 1979; Chafe, 1982; but see Poole and Field (1976) for a contrary view).
3. Preferential usage of subject-predicate constructions instead of reference-proposition (Blankenship, 1962; O'Donnell, 1974; Kroll, 1977; Ochs, 1979; Chafe, 1982).
Preferential usage of declaratives and subjunctives rather than imperatives, interrogatives, and exclamations (Portnoy, 1973; Ochs, 1979).
5. Preferential usage of passive rather than active verb voice (Blankenship, 1962; O'Donnell, 1974; Bennett, 1977; Ochs, 1979).
6. Preferential usage of definite articles rather than demonstrative modifiers and deictic terms (Ochs, 1979; Rubin, 1980).
7. Higher frequency of certain grammatical features, e.g., gerunds, participles, attributive adjectives, modal and perfective auxiliaries, etc. (Drieman, 1962; DeVito, 1966a, 1967; O'Donnell, 1974; Ochs, 1979; Chafe, 1982).
8. The need to produce complete information or idea units and make all assumptions explicit (Woolbert, 1922; Borchers, 1936; DeVito, 1965; Olson, 1977a; Goody, 1980; Rubin, 1980).
9. Reliance on a more deliberate method of organizing ideas, using such expository concepts as "thesis," "topic sentence," and "supporting evidence" (Olson, 1977a; Rubin, 1980).
10. Preferential elimination of false starts, repetitions, digressions, and other redundancies which characterize informal spontaneous speech (Woolbert, 1922; Horowitz and Newman, 1964; O'Donnell, 1974; Chafe, 1982).

Points 1–7 are syntactic features most commonly employed in written English, whereas points 8–10 are organizational and pragmatic constraints on writing. In addition to lexical and syntactic complexities, graphic signs are also used in place of prosody, context, and proxemics, to signal certain kinds of

- (a) syntactic relations:
 - lists: , ; :
 - parenthetical statements: , – ()
- (b) pauses: , ; .
- (c) illocutionary force: statement (.); question (?); exclamation (!)
- (d) emphasis: UPPER CASE, **bold face**, *italics*, underlining

In general, studies of the syntactico-semantic differences between spoken and written language attest to the non-congruence of speech and writing in “medium-related” and “message-related” dimensions (Rubin, 1980). This lack of congruence, in fact, underlies Lakoff’s recent claim that “As long as a written format is used to represent purely written discourse, there is no danger. But once we attempt to translate oral communication to the written page, we find ourselves having to translate meaning, as much as form” (1981, p. 11). This difficulty is illustrated by the problems encountered by conversational analysts working on recorded natural conversation. Discourse analysts have recognized the non-applicability of the notion of sentence to the analysis of natural conversational data, and have suggested an alternative unit of analysis under the various names of “information unit” (Halliday, 1973), “utterance chunks” (Gumperz, 1977), and “idea units” (Kroll, 1977; Chafe, 1980, 1982).

Investigators of the relationship between spoken and written language show varying degrees of awareness of the basic differences between the two modes of communication. Nearly all are, however, aware that the sentence is not a valid unit of comparative analysis of spoken and written language. Consequently, in their efforts to identify the syntactico-semantic differences between the two modes of communication, researchers have used several alternative approaches, each side-tracking the sentence as the unit of analysis.

Perhaps the earliest alternative approach is the computation of TTR. As indicated earlier, Fairbanks (1944) and Bachman-Mann (1944) initially used this method to make quantitative differentiations between “psychopathological” and “normal” language. Recently, other researchers (notably Drieman, 1962; DeVito, 1964, 1965; Horowitz and Newman, 1964; and Gibson, *et al.* 1966) have applied the same technique to the comparative analysis of spoken and written language data. However, while the computation of TTR has proved successful in accounting for the proportionate distribution of certain word classes, it has not been very useful in accounting for higher level differences (e.g., in syntactico-semantic structure) between speech and writing or in explaining why the observed differences occur.

Blankenship (1962) suggested another approach. Using a modification of Fries’s (1952) grammatical system, Blankenship proposed that the unit of analysis be the “verbal expression” defined as “any group of words functioning in relation with a verb”

(Blankenship, 1962, p. 420). Thus in the following group of words, there are four verbal expressions as indicated by the slashes:

I've just been talking at dinner with President Henry about creativity versus analysis, // and
I was tempted to say // that I wouldn't create this evening // if he wouldn't analyze //
(Blankenship, 1962, p. 420).

It can be seen from this example that Blankenship's "verbal expression" is analogous to a clause, or what is traditionally described as a "simple sentence."

When she applied her proposal to the comparison of speeches and published articles of four persons on comparable subjects, Blankenship found little variation in sentence length, some variation in sentence patterns, and some differences in the use of word classes in various positions. She also found a preponderance of transitive verbs in speech and more passive constructions in writing. In general, however, Blankenship did not recognize a clear-cut distinction between spoken and written language, and this led her to conclude that "syntactical structure is determined by an individual's style rather than by read/heard purpose" (1962, p. 422).

Blankenship's failure to find clear-cut distinctions between speech and writing can be attributed more to the nature of her data rather than to the analytical unit she chose to work with. She concentrated on formal ("planned") speeches "given before a university audience . . . and broadcast over a Radio Station" (p. 419) and formal publications of persons who are both highly articulate speakers and seasoned writers. The formality of the spoken and written samples and the linguistic sophistication of the subjects are among the variables that confounded her findings. Interestingly, in fact, Blankenship found more variation between the subjects than differences between their spoken and written samples, or between spoken and written language in general.

O'Donnell *et al.* (1967) proposed an approach which differed significantly from earlier approaches. They combined transformational concepts derived from Chomsky (1957, 1965) with the notion of "minimal terminable syntactic unit" (T-unit) proposed by Hunt (1965). Hunt's T-unit contains one independent clause and the syntactically related dependent clauses (if any). In traditional grammatical terms, a T-unit can be equivalent to a simple or complex sentence, but it is different from a compound sentence in that the latter would contain more than one T-unit (see O'Donnell *et al.*, 1967 and O'Donnell, 1974, for examples). Unlike the sentence, the T-unit is objectively identifiable both in speech and in writing. Its length reflects the number of the sentence-combining transformations applied to its underlying structure and is thus a measure of syntactic complexity (O'Donnell, 1974).

When this approach was applied to the analysis of the spoken and written samples of 30 children in grades three, five, and seven, O'Donnell *et al.* found that the average length of T-units was significantly greater in the speech than in the writing of third graders, but not of fifth and seventh graders. In the higher grades, the number of T-units became longer and more complex in written than in spoken language. Also for these grades, the number of nominal constructions, adverbial clauses, and sentence-combining transformations per T-unit was higher in the written than in the speech samples. In general, all

the measures of syntactic complexity showed significant increases with advance in grade. However, while this indicates how we can differentiate two stages of knowledge, it fails to account for why the differences occur, or how the transition from one to the other form of knowledge can be achieved (cf. Gumperz, Kaltman, and O'Connor, 1982).

In a later study employing the same approach, O'Donnell (1974) compared the responses of a highly literate individual to journalists' questions in a television program with newspaper columns written on similar issues by the same speaker and published shortly after the telecast. The contrast between the question-and-answer mode of the telecast and the highly organized and edited published account was reflected in the differences in structural complexity between the spoken and the written samples. The average length of T-units and the proportion of T-units containing dependent clauses were significantly greater in the written than in the spoken sample. Also partially accounting for these differences was the preponderance in the written samples of gerunds, participles, attributive adjectives, modal and perfective auxiliaries, and passive constructions. Speech did not, however, lag behind writing in all respects: O'Donnell also found that noun clauses, infinitives, and progressive auxiliaries were more frequent in the spoken than in the written sample. This corroborates Harrell's (1957) finding of a preponderance of noun clauses in spoken language. The validity of O'Donnell's findings is, however, constrained by one major limitation: results have been based on the language behavior of one individual. It is also not clear from O'Donnell's account how or why his written sample came to be 705 words longer than the spoken, thus negating usual expectations and previous findings (cf. Drieman, 1962). There is, however, no doubt that the T-unit approach is useful in avoiding the difficulties of objectively identifying a sentence in speech, although it is not useful in accounting for the cognitive processes underlying language production. It also fails to explain why certain grammatical units (especially clauses) are given greater prominence than others in the analysis.

Another approach, proposed by Chafe (1982), avoids these problems by identifying "idea units" as the units of analysis. Derived from observations of spontaneous speech, idea units usually exhibit one of a small set of syntactic structures, are typically characterized by a coherent unit of intonation contour, and are bounded by pauses. Each idea unit has a mean length of approximately 2 seconds or approximately six words each (Chafe, 1979, 1980). The notion of idea units is based on the "spurt-like" nature of speech which probably reflects the "jerky" nature of thought (Chafe, 1982). Chafe speculates that if language reflects the pace of thought, then language production is carried out at the rate of production of idea units both when overtly vocalizing or when thinking to oneself while writing.

Hypothesizing that writing will exhibit a higher degree of "integration" than speech, Chafe compared the structure of idea units in the speech and writing of 14 subjects (faculty and graduate students). The spoken samples were based on informal dinner-table conversations and the written samples on written academic papers by the same subjects. Counting the distribution of certain recurrent word classes and grammatical structures in the samples, Chafe found that formal written language differs from informal speech by having a larger proportion of nominalizations, attributive adjectives, genitive subjects and objects, participles, conjoined phrases, series, sequences of prepositional phrases,

complement clauses, and relative clauses. However, it could be objected that the spoken and written samples are not on comparable topics. In addition, differences in setting, context, and purpose between formal (published) academic papers and informal (dinner-table) conversations are in themselves significant variables that may affect lexical and syntactic choices. Consequently, the findings seem to result more from the maximization of contrasts in the data base than from differences in modality.

Using the concept of "idea unit," Tannen (1980) avoided the noncomparability of topic by comparing the spoken and written versions of a narrative given by the same subject. She concluded from her findings that written narratives often make use of certain features of spoken language such as informal registers, direct quotations, and details. While Tannen's data are consistent with her findings, other possibilities must not be overlooked. Depending on the method of data collection and the purpose of narration, written personal narratives can approximate either speech or writing. If "unplanned" and consequently less structured than, say, written academic papers, written personal narratives can exhibit many features of spoken language. However, if it is "planned," written personal narratives will exhibit more of the features of written language (Ochs, 1979). This, no doubt, explains why Ochs's comparison of spoken and written narratives produced results that contrast with Tannen's, namely that written narratives manifest linguistic features that are distinctively different from those of spoken language. Similarly, in his comparison of oral and written narratives of young school children, Harrell (1957) found clear structural differences between the two modes of communication. In particular, he found significant variations in the distribution of subordinate clauses. While noun clauses were more frequent in speech, adverbial, adjectival, and relative clauses were more frequent in the written samples.

Although Bernstein himself never discusses the differences between spoken and written language at all, it has been suggested by exponents of his model that his distinction between restricted and elaborated codes parallels the distinction between spoken and written language respectively and is therefore related to the acquisition of literacy (Stubbs, 1980, p. 111). Thus, as indicated in the introduction, Bernstein's code elaboration model stimulated a number of studies of the relationship between spoken and written language (e.g., Lawton, 1963; Greenfield, 1972; Poole and Field, 1972, 1976). For example, building on previous research by Harrell (1957), Drieman (1962), DeVito (1966a), and Golub (1969), Poole and Field (1976) compared oral and written language in terms of a Bernstein code elaboration model (Bernstein, 1961, 1964; Lawton, 1963). Their data were based on samples obtained from freshman university students and their focus was on the distribution of major lexical and syntactic categories (adjectives, adverbs, personal pronouns, subordinate clauses, verb groups, preverb structures, etc.). Their results indicate that, although spoken language is more complex than written language in the area of syntax, written language demonstrates more overall structural complexity than spoken language. For example, they found that written language contains more adjectival elaboration, more complex verb groups, and fewer indices of personal reference. They also found that, in general, the same differences between spoken and written language characterize both middle- and working-class groups in their samples, implying that the problems of transition from spoken to written

language are universal for all children (Olson, 1977b; Cook-Gumperz and Gumperz, 1981).

One significant aspect of Poole and Field's findings which contrasts sharply with popular views is the distribution of subordinate clauses. They found that more subordinate clauses were used in speaking than in writing, whereas several studies have found the reverse to be true both for children (Harrell, 1957) and for adult subjects (O'Donnell, 1974; Kroll, 1977; Chafe, 1982). This seeming oddity in Poole and Field's findings may have been due to the highly structured nature of the oral data. Their speech samples were obtained individually from freshman students (who probably took the experiment very seriously), using the Lawton-type structured interview schedule containing questions designed "to tap various levels of linguistic coding complexity, e.g. description to abstraction" (Poole and Field, 1976, p. 307). The resultant data were highly structured and, in certain areas, more complex than the written data.

But perhaps the most divergent results were those of Horowitz and Newman (1964). By controlling topical and contextual variables, and limiting time both for preparation and for exposition, Horowitz and Newman found that spoken expression produced "more ideas and subordinate ideas," in addition to being "more productive and prolific" than written expression. It should be noted however, that in contrast to Chafe's concept of "idea units," Horowitz and Newman use the notion of idea unit to refer to a *cognitive* rather than a syntactic category:

An idea is an utterance that expresses a thought in a meaningful, relevant, and unique way. By meaningful we mean an exposition of material that makes sense, cognitively, to the experimenter . . . in expressing the thought, or concept, there need not be a subject and predicate, nor need there be grammatical correctness (1964, p. 642).

Their further distinctions between major, subordinate, and ancillary ideas (see p. 642, for examples) are equally non-syntactic. They were, therefore, not making a syntactic argument, as has often been cited in the literature (e.g., by Poole and Field, 1976), when they declared: "When ideas and subordinate ideas are combined (I + SI), spoken expression is more productive than written expression" (Horowitz and Newman, p. 644). Their hypotheses and experimental design favored spoken language as did the reasons they advanced for the differences: (1) speaking is mechanically and psychologically easier than writing; (2) there is the tendency in speaking to avoid silence. When equal preparation and exposition time was given for the spoken and written samples, Horowitz and Newman found that the above features facilitated the production of "more ideas" in spoken than in written language. Even when more facile writing techniques (e.g., typing and stenotyping) were used, it was found that "spoken expression produces significantly more cognitive and linguistic material than written expression" (Horowitz and Berkowitz, 1964, p. 617).

The above review shows that research on the lexical and syntactico-semantic differences between spoken and written language have yet to overcome a number of problems in order to validate their findings. First is the need for data control. More often than not, conclusions about the differences between spoken and written language are an artifact

of data choice. For example, when data were based on the formal speeches and published writings of the same individuals, there was no significant difference in sentence length (Borchers, 1936; Blankenship, 1962) or in comprehensibility (DeVito, 1964, 1966a). Similarly, when informal spoken narratives were compared with their informal written version, it was difficult to differentiate the two modes (Tannen, 1980). However, when the informal and formal poles were kept distinctively apart (e.g., informal dinnertable conversations vs. formal academic papers), significant structural differences emerged between spoken and written language (O'Donnell, 1974; Chafe, 1982). The question thus arises as to whether comparison is between spokenness and writtenness or between informality and formality in discourse.

A second problem is lack of adequate definition of the variables that may affect the findings. For example, Poole and Field (1976) and Kroll (1977) produced opposite results comparing use of subordination in oral and written language despite the common assumptions they shared (1) that writing is more complex than speech and (2) that writing is "planned" whereas speech is "unplanned" (Ochs, 1979). Poole and Field found more subordination in their oral than in the written samples, while Kroll's findings showed the reverse to be true. However, a close comparison of the two studies shows that they differ significantly in data base, especially in the topics on which the spoken and written samples were based, in the communicative tasks in which subjects were involved, and in the method of data collection and analysis. Besides, both researchers failed to specify what kind of planning and pragmatic constraints were involved in their spoken and written samples and how these affected their results.

Thirdly, like many studies of the consequences of literacy, comparative studies of spoken and written language have emphasized general, rather than specific, consequences of writing on language structure, the working assumption being that written language is generally more complex than spoken language. Yet contradictions occur when different studies are compared on specific findings. Failure to compare speech and writing from point of view of how specific communicative tasks are achieved in either modality is partly responsible for the contradictions in the literature. Thus studies of differences between spoken and written language would do better to begin with a comparison of the devices used in both modalities to perform the communicative tasks of narrating, describing, explaining, arguing, denying, protesting, emphasizing, etc. (see Gumperz, Kaltman, and O'Connor, 1982).

A fourth problem with studies of spoken and written language is their quantitative orientation, each researcher deciding on what and how to count. It is commonplace in social science that statistical counts usually capture only "etic" rather than "emic" categories, thereby ignoring the underlying logic behind surface behavior. That researchers are not even agreed on statistical differences between spoken and written language shows that the surface phenomena have yet to be fully understood, let alone the underlying logic. This partly explains why researchers often give obvious reasons as explanations of the structural differences between spoken and written language (see Section 4 below).

A fifth problem, previously noted in Section 2, is the limitation of the samples to university students and professors, ignoring the rest of the population and the various

uses of speaking and writing. Perhaps the best solution to this parochialism is to conduct ethnographic studies of speech and writing activities in different communities within a given population, taking samples from elementary, high school, and university communities; from bureaucratic, social, and religious organizations; from the press and various work places; and from ethnographic "diaries" of daily uses of speech and writing, taking note of differences in age, sex, social class, cultural background, and the like. Such ethnographic study will surely remove research on the relationship between speech and writing from the shackles of (laboratory) experimental orientations. In conducting such ethnographic studies, more attention should be given to comparisons of formal speech vs. formal writing and of informal speech vs. informal writing.

This leads us to the sixth shortcoming of contemporary comparisons of speech and writing: neglect of oral ritual language (e.g., ritual chants and divination texts) which exhibits the formal, functional, and grammatical features of written language (Akinnaso, 1981a, 1981b, 1981c). Recent research has shown that oral ritual and written language are both acquired in formal settings; they are both used for storage and transmission of knowledge and cultural traditions; they both derive from a transcendental source and are equally cited to validate social action; and they make use of similar processes of lexical elaboration and syntactico-semantic complexities (see especially Akinnaso, 1981a, 1981c; Olson, 1980). These far-reaching similarities between oral ritual and written language have repercussions for the "linguistic discontinuity hypothesis" and for studies of language specialization (see Akinnaso, 1981a for more detailed discussion of these issues).

4. WHY DIFFERENT?

Researchers have advanced various reasons for the structural differences between spoken and written language (see, e.g., Horowitz and Newman, 1964; Horowitz and Berkowitz, 1964; DeVito, 1966b; Chafe, 1979, 1982). In general, it is argued that spoken and written language are structurally different because they differ in their modes of acquisition; in their methods of production, transmission, and reception; and in the ways in which elements of structure are organized (or in degree of "planning").

Speech is normally acquired naturally without formal instructions (in family settings, on playgrounds, on the street, etc.), whereas writing has to be consciously learned, usually in the formal setting of the school. Such a setting is often more authoritarian and heavily reliant on standardized learning activities and systematized procedures. All along in the process of acquiring writing skills, the learner is taught to pay particular attention to the choice of words and their arrangement, appealing to such notions as "grammaticality," "correctness," and "proper organization". This emphasis on prescriptivism is further enhanced by the numerous regulations imposed on writing by school teachers, examining bodies, bureaucrats, journal editors, publishers, and professional writing manuals.

The formality surrounding the acquisition and use of writing makes it a more deliberate activity than speech. Writing, says Ong, depends on "consciously contrived

rules" and "absent audiences" (1980, p. 200). Not only is the audience usually absent, the writer must be isolated for the optimization of his/her performance. Ong's observations tie in with the basic psychological claim that writing requires conscious analytical process, being acquired and used *out of context* rather than *in context* (Vygotsky, 1962; Bruner, 1968; Bartlett and Scribner, 1977). What this means is that speech presupposes a common context and point of view for both speaker and listener, whereas writing is relatively independent of context, being addressed to no one in particular, or, at best, to a distant interlocutor (Goody and Watt, 1963; Greenfield, 1972). Goody and Watt (1963, p. 321) summarize the difference thus:

Writing established a different kind of relationship between the word and its referent, a relationship that is more general and more abstract, and less closely connected with the particularities of person, place, and time that obtains in oral communication.

This statement has two interrelated implications for writing: explicitness and decontextualization (Kay, 1977; Olson, 1977a, 1977b; Akinnaso, 1981b). By decontextualization is meant the abstraction of a statement from temporal, spatial, and situational boundaries in such a way that the signalling of intention becomes concentrated in the linguistic channel. This reliance on a single modality for the communication of intent has been aptly characterized by Kay (1977) and Olson (1977a, 1977b) as language autonomy. With increasing explicitness, Olson argues, written language becomes more able "to stand as an unambiguous or autonomous representation of meaning" (1977a, p. 258).

While writing is uni-modal, speech is multi-modal, making use of linguistic, prosodic, kinesic, and contextual cues in the signalling of meaning. Consequently, spoken language not only expresses propositional, emotional, contextual, and culturally specific messages, but also signals illocutionary force. Written language, on the other hand, expresses mainly propositional messages, being minimally dependent on a common frame of reference between sender and receiver of messages, in addition to lacking necessary non-propositional signalling devices (Greenfield, 1972; Vachek, 1976; Olson, 1977b). As indicated earlier, only three basic types of illocutionary force can be signalled in writing without the need to lexicalize what is intended. Similarly, attempts to convey prosodic and contextual information in writing often lead to lexical elaboration and syntactic complexity (Gumperz, Kaltman, and O'Connor, 1982 — to be discussed later).

The multi-modality of speech results partly from its spatial, temporal, and cultural boundedness and partly from the medium through which speech is manifested. Speech is phonically manifested, transmitted as sound waves, and so received by the hearer, whereas writing is graphically manifested, transmitted by light waves, and perceived through the eyes. Because of the physical properties of speech sounds, spoken language fades rapidly, and it is prototypically dependent on the co-presence or proximity of the sender and receiver of messages. On the other hand, because of the permanency and transportability of written language, writing is the medium prototypically used between interlocutors that are separated in time and space. Although recent research has shown that speech and writing share similar phonological and semantic processing (Posner and

Hanson, 1980), it is equally evident that there are several aspects unique to one modality (Hanson, 1981). For example, while kinesic and other non-linguistic variations (e.g., pitch and intonational contrasts) can affect the semantic content of spoken messages considerably, such variations are not easily expressible in writing. Thus writing depends on much fewer modality-specific contrasts than speech for the communication of messages (cf. Francis, 1958, p. 339; DeVito, 1966b, p. 57; Vachek, 1976).

Some modality-specific differences between speech and writing arise from differences in the situations in which the two modes of communication are used. In many situations, speech and writing are in complementary distribution. For example, natural conversations are always carried out in spoken language, whereas, in modern industrial societies, speech is inappropriate for much bureaucratic communication such as applying for a job, requesting social services, filling out tax and credit application forms, and so on. Similarly, spoken language is not sufficient for the storage and transmission of information in industrially complex and ethnically diverse speech communities such as the United States. These functional and situational differences have lexical, syntactic, and semantico-pragmatic implications for the two modalities. For example, speech abounds in several constructions concerned with the mechanics of inter-personal relations such as self-reference words, tag-questions, and commands which rarely occur in writing. On the other hand, there are several lexico-syntactic and graphic devices that are predominantly or exclusively used in written English as devices for expressing meanings specifically through the written mode — devices which may, in fact, not be acquired at all until the speaker learns the conventions of written expression (Gumperz, Kaltman, and O'Connor, 1982).

Speech and writing also contrast in manner and speed of production. Writing is essentially a mechanical process, requiring the manipulation of a physical tool and the conscious coordination of specific motor and cognitive skills. Thus writing is “completely and irremediably artificial” (Ong, 1980, p. 199), whereas speech is a natural process, making use of so-called “speech organs.” Writing and written language are not universal phenomena, whereas speech is a universal, being crucial to the human condition (Goody, 1977, 1980). While the naturalness of speech aids its speed of production, the mechanical process of writing slows down its speed of production. Chafe (1982) puts the average speed of spoken English in the range of 180 words per minute, while contending that writing typically takes place at slower than one-tenth the speed of speaking. However, DeVito (1966b) thinks that the difference is not that much, putting writing at about one-fifth the speed of speaking. Such disagreements on details no doubt result from basic differences in researchers’ data base. In any case, what is important here is that writing is much slower than speaking.

Some have picked on this contrast to contend that the slowness of writing favors the faster deployment of cognitive processes in such a way that our thoughts must constantly get ahead of our expression, resulting in the availability of time to integrate a succession of ideas into a single linguistic whole in a way that is not available in speaking where the expression of thought is more or less simultaneous with its formulation (Horowitz and Newman, 1964; DeVito, 1966b; Chafe, 1982). Chafe (1982) summarizes the argument thus:

... the abnormal quickness of reading fits together with the abnormal slowness of writing to foster a kind of language in which ideas are combined to form more complex idea units and sentences. I will say that written language tends to have an "integrated" quality which contrasts with the "fragmented" quality of spoken language (p. 6).

However, while the slowness of writing may affect its organization, it cannot be the only or major factor accountable for the integrated quality of written language. Equally, if not more, important are the features of permanency, surveyability, and (re)organization which characterize written language. These features make it possible to (re)organize both thought and expression in a way that is not possible in speech. We shall discuss each of these features in turn.

Unlike speech which fades rapidly, unless recorded, writing is permanent and reproducible. This, according to Martinet, is the principal cause of difference between speech and writing:

The chief reason why people do not write as they speak is probably that since writing leaves permanent traces, whereas speech unless recorded is lost forever, writers are far more careful than speakers . . . Consequently, written style is not spoken style (1962, pp. 122-123).

Moreover, the permanency of writing and the opportunities of visual inspection make it possible to review the entire text of what has been written, making relevant changes as desired. Writing affords the linear representation of thought in visuo-spatial form; its permanency and surveyability facilitate rethinking and reorganization of expression: "words are laid out clinically on the page . . . , capable of being struck out, re-ordered, substituted, pored over, reflected upon" (Goody, 1977, p. 157). Thus the increased formalization and the high integration of written language result not so much from the slowness of writing as from the visuo-spatial characteristic of writing and from the subsequent opportunity to scrutinize and reorganize discourse in many different ways.

Organization, of course, inevitably involves many deliberate activities, the most important of which is planning. Writing is not only planned but also plannable, whereas speech, in the form of spontaneous conversation, is unplanned and unplannable (Ochs, 1979). Ochs suggests that for discourse to be unplanned and unplannable, it must lack forethought and prior organization and it must be impossible to predict its form and content well in advance. On the other hand, planned and plannable discourse is characterized by forethought, prior organization, and a predictable message content. She goes further to hypothesize four features that characterize and distinguish unplanned and planned discourse of (educated) English speakers:

1. In unplanned discourse more than in planned discourse, speakers rely on immediate context to express propositions.
2. In unplanned discourse, speakers rely on morpho-syntactic structures acquired in the early stages of language development, whereas planned discourse makes greater use of morpho-syntactic structures that are relatively late to emerge in language.
3. In unplanned more than in planned discourse, speakers tend to repeat and replace

lexical items in the expression of a proposition.

4. In unplanned discourse, the form and content of sequentially arranged social acts tend to be more similar than in relatively planned discourse.

Drawing partly from data obtained from the spoken and written versions of narratives delivered by the same speakers, Ochs argued that the syntactico-semantic differences between spoken and written language result from these fundamental distinguishing characteristics of unplanned and planned discourse.

Interestingly, the distinctions Ochs draws between unplanned and planned discourse parallel those drawn by Bernstein between restricted and elaborated codes. Thus unplanned discourse corresponds to restricted code while planned discourse parallels elaborated code. Furthermore, both Ochs and Bernstein agree that lexical and syntactic predictability is more probable in unplanned discourse/restricted code than in planned discourse/elaborated code largely because the range of available lexical and syntactic choices is more restricted in the former. When Poole (1972, 1973) tested this hypothesis, using the cloze procedure and an information theory approach, she found that restricted code (used by working-class subjects) is, indeed, more predictable in lexicon and syntax than elaborated code (used by middle-class subjects). To the extent that restricted and elaborated codes can be respectively equated with spoken and written language, it can also be argued that written language would be characterized by less predictable lexicon and syntactic structure than spoken language. Availability of a relatively more extensive source of lexical and syntactic alternatives and the opportunity to consult various sources (dictionaries, experts, etc.) are among the factors that mitigate lexico-syntactic predictability in writing. It should be noted, however, that lexical and syntactic predictability is mediated by other variables than modality-specific features. Such intervening variables include register peculiarities, speakers' social class, and level of linguistic knowledge (cf. Poole, 1972).

Although Ochs recognizes predictability as a dimension of contrast between unplanned (spoken) and planned (written) discourse, her major emphasis is on planning. However, while recognizing emphasis on planning as a major characteristic of writing, it should be noted that much of the refinement of written language (especially academic writing) often takes place in the course of successive revisions of preliminary drafts, usually after suggestions by colleagues and/or editors have been incorporated. If we examined corrected, deleted, or substituted items in preliminary drafts, we would discover that hesitations, false starts, and several imperfections that are usually associated with speech abound also in writing. However, such imperfections are edited out so that the finished product often bears no perceptible signs of them. Among the factors which make this possible are the permanency and surveyability of writing, factors which facilitate the monitoring of cognitive and linguistic processes as well as the editing of recorded material. Speech cannot be so edited because of spatio-temporal limitation.

5. TOWARD A NEW APPROACH

While the above explanations help us to know why spoken and written language employ different surface features, they still fail to explain why certain features, and not others, are preferred in a particular modality. Secondly, we still need to know how writers cope with the shortage of devices for signalling specific messages or semantic contrasts. For example, we want to know how information signalled through prosodic, kinesic, and contextual cues in speech are signalled in the written mode. Finally, we want to go beyond mere explanations of surface differences to consideration of the theoretical and pedagogical implications of the transition from speech to writing.

Recent work by Gumperz and his students provides a theoretical and methodological approach to the study of differences between spoken and written language which goes beyond polar comparisons and taxonomic differentiations (Cook-Gumperz and Gumperz, 1981; Gumperz, Kaltman, and O'Connor, 1982; Collins and Michaels, 1980; Michaels, 1981a). Their investigation of the differences between spoken and written language is carried out within the broader context of the acquisition of literacy and within the framework of conversational analysis being developed by Gumperz (Gumperz, 1976, 1977, 1978, 1980, 1982). In these studies, Gumperz has demonstrated that participants in a conversation draw on culturally specific "discourse strategies" in performing specific "communicative tasks" at different points in the conversation. Discourse strategies include linguistic, prosodic, and proxemic cues. These "contextualization cues" are used by conversationalists to "signal interpretative frames" (Gumperz, 1977, p. 199) and to achieve conversational cooperation. Central to this framework of conversational analysis is the notion of "thematic cohesion," i.e., how elements in a text are unified (or disconnected from each other) in such a way as to signal a continuously developing theme (cf. Halliday and Hasan, 1976).

Since thematic cohesion is central to the organization of spoken and written texts, it is proposed that differences between speech and writing can be better explained through investigation of how thematic cohesion is achieved in either modality (Gumperz, Kaltman, and O'Connor, 1982). The basic question that Gumperz and his collaborators seem to be dealing with in their comparison of spoken and written language is: What are the linguistic consequences of the written mode (and generally of literacy) and what linguistic and cognitive problems does the student face in learning the relevant options of the written mode (or in trying to acquire literacy)? One of the major problems they identified is how to accomplish the shift in modality to the discourse strategies, especially the cohesive devices, of the written mode. Daniel Jones had earlier recognized this problem, but he did not investigate it; he declared:

Most of the shades of meaning that speech is capable of conveying by the above means [stress, intonation, etc.] are either incapable of expression in writing or can be expressed only by added words or by alterations of wording (1943, p. 207).

While Cook-Gumperz and Gumperz (1981) provided detailed explanations of the general linguistic and cognitive problems confronting beginning learners in the transition

to literacy, Gumperz, Kaltman, and O'Connor (1982) compared spoken and written discourse, using literate, neo-literate, and preliterate subjects. After a review of the literature, they concluded that the observed surface differences between spoken and written language can be attributed to modality-specific pragmatic constraints such as: (1) availability or non-availability of feedback and of knowledge of one's audience and their presuppositions; and (2) availability or non-availability of reference to preceding material and differential reliance on memory (p. 26). They went further to contend that these constraints have differential effects on the cohesive devices used in speech and in writing, and are therefore accountable for the modality-specific differences in lexico-syntactic choice. Through systematic analyses and comparisons of illustrative data on specific communicative tasks, they showed how the transition from speech to writing is largely a matter of specifying through lexicalization, syntactic complexity, and punctuation material which was signalled through prosodic, contextual, and other nonlinguistic cues in speech. Their findings indicate that writers must learn to shift signalling levels in order to achieve the relevant cohesive devices of written language, suggesting that "the problems that occur in the acquisition of literacy, . . . especially in the production of . . . expository prose . . . are thus in some sense similar to those which occur in communicating across ethnic boundaries" (p. 4).

It should be noted that Gumperz, Kaltman, and O'Connor did not intend a dichotomy by the above statement since they admit that speech and writing cannot be regarded as two independent codes, regardless of superficial differences between them, because of the mutual influences they continue to exert on each other (pp. 23-24). The mutual dependency of speech and writing has historical and class dimensions. In nearly all communities where writing is employed as a means of wider communication, the written style is often based on the speech patterns of the elite, ruling, or dominant group. In speech communities that have had a long history of literacy (e.g., England and the United States), a situation develops whereby (many) members of the dominant group begin to "speak a written language" (Greenfield, 1972; Olson, 1977b) or use a "literate speech style" (Collins and Michaels, 1980). This, in part, underlies the working assumption of many psychologists, discourse analysts, students of literacy, and sociologists of education that children who, by their upbringing, are exposed to "literate" discourse strategies have a greater advantage in acquiring literacy than children that are not so exposed (Bernstein, 1964, 1971; Bruner, 1968; Greenfield, 1972; Olson, 1977a, 1977b; Cook-Gumperz and Gumperz, 1981; Collins and Michaels, 1980; Wells, 1981; Gumperz, Kaltman, and O'Connor, 1982; Michaels, 1981a, 1981b).

Thus the assumption that Western spoken tradition is itself a literate one underlies the comparative study of conversational styles by Collins and Michaels (1980). Using the concept of thematic cohesion and distinguishing heuristically between what they call "literate" and "oral" style, they found that the former favors the use of a wide variety of lexical and syntactic devices, especially noun complements, to signal agent focus, causal connections, coreference relations, and so on, whereas "oral" style relies more on prosodic cues, especially distinct contours, to signal similar relations and distinctions. An interesting aspect of this study is the preponderant use of nominal structures by speakers of the "literate" style. Halliday (1967) had earlier made a similar point about the

increasing reliance on the "nominal style" by speakers and writers of British English. He explained this phenomenon in terms of the historical, technological, and literary development of the West. While Hawkins (1969) and Henderson (1970) have demonstrated the social-class correlates of the nominal style, Collins and Michaels highlighted the pedagogical implications of this discourse style by pointing out how negative evaluation in the classroom can result from failure to use it.

6. CONCLUSION

While granting that the question of the relationship between spoken and written language can hardly be settled in a single paper, the present account would be incomplete without exploring how speech and writing relate to prototypical forms of language and meaning. There are two contradictory theoretical positions on the matter. There is the traditional view which stresses the dependence of written language on speech, the assumption being that phylo- and ontogenetically written language was acquired by means of spoken language. In this view, reading and writing are considered phonemically mediated performances, both directly linked to meaning (Bloomfield, 1933; Rubenstein, Lewis, and Rubenstein, 1971). An alternative position emphasizes the progressive autonomy of the graphemic system during the development of reading and writing skills, the assumption being that meaning may be directly extracted from visual input or linked to written output without phonological mediation (Forster and Chambers, 1973). Friederici, Schoenle, and Goodglass (1981) offer a review of both positions and report experiments validating both phonologically-mediated and direct-pathway decoding of written language. While the traditional view stresses that speech and writing are grammatically the same with only slight lexical differences, the "autonomy" view implies that speech and writing can have quite different surface forms, while sharing the same underlying semantic system.

Vygotsky's view on the subject is probably closer to the "autonomy" view. He argued in *Thought and Language* (1962) that although the development of writing ability is predicated on the internalization of language as "inner speech", the potentialities of written language go far beyond the powers of speech:

[In spoken language], every sentence is prompted by a motive. Desire or need lead to request, question to answer, bewilderment to explanation. The changing motives of the interlocutors determine at every moment the turn oral speech will take. It does not have to be consciously directed — the dynamic situation takes care of that. The motives for writing are more abstract, more intellectualized, further removed from immediate needs. In written speech, we are obliged to create the situation, to represent it to ourselves (1962, p. 99).

Here Vygotsky paints a situation in which writing is linked with higher cognitive processes and cannot, therefore, be viewed as inferior to speech, implying that spoken language is necessary but not sufficient for performing higher cognitive functions. In performing these elevated functions, he argued, written language can take a form of its own.

Somewhat less optimistic are Prague School linguists who stress that speech and writing are stylistic and functional variations of the same linguistic system (Vachek, 1948, 1976). They are, however, not as pessimistic as traditional American structuralists concerning the status of written language. Responding to the traditional view that writing is inferior to speech, Vachek argues:

The development of a community's higher culture and civilization is unquestionably conditioned by the existence in its language of a written norm, *the* vehicle of higher needs and wants of the community. It would, then, be completely out-of-place to brand the written norm as an inferior kind of structure (1976, p. 415).

Like Vachek, Smith endorses the functional and stylistic differences between spoken and written language: "It should not be surprising or anomalous that differences exist between spoken and written language; they are generally used for quite different purposes and addressed to quite different audiences" (1978, p. 80). Like Vygotsky and Bruner, Smith contends that these differences have cognitive dimensions. To acquire literacy, he says, two fundamental cognitive insights are needed: (1) that print is meaningful, and (2) that written language is different from speech (Smith, 1977). It should be noted, however, that Smith is not polarizing spoken and written language: "Speech and print are not different *languages* — they share a common vocabulary and the same grammatical forms — but they are likely to contain different distributions of each" (1978, p. 80).

This paper has been concerned largely with the nature of these distributions, with *how* and *why* spoken and written language differ. A review of the evidence supports the view that spoken and written language derive from the same semantic base, making use of the same lexico-syntactic system, and varying mainly in the choice and distribution of vocabulary and syntactic patterns in response to modality-specific pragmatic constraints. A measure of this variation is to be found in strategies employed to achieve thematic cohesion in either modality (Gumperz, Kaltman, and O'Connor, 1982). Thus while speakers often automatically employ certain prosodic or kinesic cues to signal particular intentions, writers would worry about which words and/or syntactic patterns would effectively signal the same intentions.

What this implies is that differences between spoken and written language should be investigated within the broader framework of discourse management rather than within the narrower framework of quantitative differentiation of lexical and syntactic tokens. Furthermore, while factual and scientific knowledge of differences between spoken and written language is interesting, the theoretical and pedagogical implications of such differences are undoubtedly more interesting. The strength of the new approach reviewed in Section 5 lies partly in its clarification of major theoretical issues and partly in its pedagogical utilities. Thus current work by Gumperz and his collaborators promises us concrete data and findings that could be used in preparing instructional material on language for beginning literacy programs.

This review highlighted the need to reframe or raise new questions on language research. Such new formulations would include the following fundamental questions:

1. How are the basic functions of writing (especially information storage and transmission) performed in nonliterate communities?
2. What are the consequences of oral composition and oral ritual communication for language structure and communicative processes?
3. How do speech and writing relate in tonal languages (e.g., Yoruba) where intention and semantico-pragmatic contrasts are signalled in speech mainly through lexical and syntactic choice rather than prosody (intonation, stress, pitch, etc.)?
4. How do speaking and writing abilities vary with age, sex, social class, cultural background, and the like, within and across speech communities, and from generation to generation?
5. What are the cognitive processes involved in learning to write in one's native language or dialect? How do they compare with the processes underlying the acquisition of writing in another language?
6. Which areas of spoken language are most affected by the written form in societies that have had a long history of literacy?
7. What is the role of written language within the educational system in contemporary societies?
8. What changes in educational policy and curricula are necessary to teach the language features peculiar to writing?

Research on the first two questions is particularly urgent in view of the current claims by some educators, psychologists, and linguists regarding the assumed supremacy of "written language" and "literate culture" vis-à-vis the impoverishment of "oral language" and "oral culture". The direction of such research has been indicated elsewhere (Akinnsaso, 1981a, 1981b, 1981c). The new Cambridge series on studies of oral and literate culture edited by Peter Burke and Ruth Finnegan promises us further data and findings that could be used as a basis for comparing orality and literacy across time and space.

While most of the distinctions made between spoken and written language in this paper may apply to English, there is no reason to believe that they will apply to all languages in general, partly because of structural and historical differences between languages and partly because the history, development, and functions of writing vary from one linguistic community to another. For example, current research on oral and written Yoruba shows that oral ritual Yoruba (e.g., divination texts) is sometimes more "elevated" than other varieties, spoken or written (Akinnsaso, 1981a, 1981c). Research based on question 3 would therefore be a test case for the approach proposed by Gumperz, Kaltman, and O'Connor (1982).

Question 4 should be a major concern of ethnographic studies of speech and writing as suggested in Section 3 while attention to question 5 will enhance the pedagogical relevance of language acquisition studies. Attempts at language education reforms can hardly succeed without due attention to questions 4–8; yet they have been largely ignored in most investigations. But perhaps the major importance of these questions is to

direct comparative studies of speech and writing toward interesting theoretical and pedagogical problems and thus wean them from bondage to quantitative differentiation of tokens.

REFERENCES

- AKINNASO, F.N. (1981a). The literate writes and the nonliterate chants: Written language and ritual communication in sociolinguistic perspective. Paper presented at the Third Delaware Symposium on Language Studies, Newark, Delaware.
- AKINNASO, F.N. (1981b). The consequences of literacy in pragmatic and theoretical perspectives. *Anthropology and Education Quarterly*, **12**, 163-200.
- AKINNASO, F.N. (1981c). Elevated language in nonliterate societies. Paper presented at the 25th Annual Meeting of the Kroeber Anthropological Society in collaboration with the Berkeley Sociolinguistics Group, Berkeley, California.
- ANDERSON, T.H., ARMBRUSTER, B.B. and KANTOR, R.N. (1980). How clearly written are children's textbooks? *Reading Education Report No. 16* (Center for the Study of Reading, Urbana Champaign).
- BACHMAN-MANN, M. (1944). The quantitative differentiation of samples of written language. *Psychological Monographs*, **56**, 41-74.
- BARTLETT, E. and SCRIBNER, S. (1977). In reference to writing: Some analyses of texts produced by children. Paper presented at the National Institute of Education Conference on Writing.
- BENNETT, T. (1977). Verb voice in unplanned and planned narratives. In E.O. Keenan and T. Bennett, (eds.), *Discourse Across Time and Space. Southern California Occasional Papers in Linguistics*, **5**, (Los Angeles).
- BERNSTEIN, B. (1961). Aspects of language and learning in the genesis of the social process. *Journal of Child Psychology and Psychiatry*, **1**, 313-324.
- BERNSTEIN, B. (1964). Elaborated and restricted codes: Their social origins and some consequences. In J.J. Gumperz and D. Hymes, (eds.), *The Ethnography of Communication*. (Special Issue of American Anthropologist **66**(6), Part 2), pp. 55-69.
- BERNSTEIN, B. (1971). *Class, Codes and Control, Vol. 1: Theoretical Studies towards a Sociology of Language* (London).
- BLANKENSHIP, J. (1962). A linguistic analysis of oral and written style. *Quarterly Journal of Speech*, **48**, 419-422.
- BLOOMFIELD, L. (1933). *Language* (New York).
- BORCHERS, G. (1936). An approach to the problem of oral style. *The Quarterly Journal of Speech*, **22**, 114-117.
- BROWN, R. and BELLUGI, U. (1964). Three processes in the child's acquisition of syntax. *Harvard Educational Review*, **34**, 133-151.
- BRUNER, J.S. (1968). *Processes of Cognitive Growth* (Worcester, Mass.).
- CHAFE, W.L. (1979). Integration and involvement in spoken and written language. Paper presented at the 2nd Congress of the International Association for Semiotic Studies, Vienna.
- CHAFE, W.L. (1980). The deployment of consciousness in the production of a narrative. In W.L. Chafe, (ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production* (Norwood, N.J.), pp. 9-50.
- CHAFE, W.L. (1982). Integration and involvement in speaking, writing, and oral literature. In D. Tannen, (ed.), *Oral and Written Discourse* (Norwood, N.J.), in press.
- CHOMSKY, N. (1957). *Syntactic Structures* (The Hague).
- CHOMSKY, N. (1965). *Aspects of the Theory of Syntax* (Cambridge, Mass.).
- COLLINS, J. and MICHAELS, S. (1980). The importance of conversational discourse strategies in the acquisition of literacy. *Proceedings of the 6th Annual Meeting of the Berkeley Linguistics Society*, pp. 143-156.

- COOK-GUMPERZ, J. and GUMPERZ, J.J. (1981). From oral to written culture: The transition to literacy. In M.F. Whitehead (ed.), *Writing: The Nature, Development, and Teaching of Written Communication*. Vol. 1, *Variation in Writing: Functional and Linguistic-Cultural Differences* (Hillsdale, N.J.), pp. 89-109.
- DEVITO, J.A. (1964). A quantitative analysis of comprehension factors in samples of oral and written technical discourse of skilled communicators. Unpublished Ph.D. thesis, University of Illinois.
- DEVITO, J.A. (1965). Comprehension factors in oral and written discourse of skilled communicators. *Speech Monographs*, **32**, 124-128.
- DEVITO, J.A. (1966a). Psychogrammatical factors in oral and written discourse by skilled communicators. *Speech Monographs*, **33**, 73-76.
- DEVITO, J.A. (1966b). The encoding of speech and writing. *The Speech Teacher*, **15**, 55-60.
- DEVITO, J.A. (1967). Levels of abstraction in spoken and written language. *Journal of Communication*, **17**, 354-361.
- DEWEY, G. (1923). *The Relative Frequency of English Speech Sounds* (Cambridge, Mass.).
- DOWNING, P. (1980). Factors influencing lexical choice in narrative. In W.L. Chafe, (ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production* (Norwood, N.J.), pp. 89-126.
- DRIEMAN, G.H.J. (1962). Differences between written and spoken language. *Acta Psychologica*, **20**, 36-57; 78-100.
- FAIRBANKS, H. (1944). The quantitative differentiation of samples of spoken language. *Psychological Monographs*, **56**, 19-38.
- FORSTER, K.I. and CHAMBERS, S.M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, **12**, 627-635.
- FOSSUM, E.C. (1944). Analysis of the dynamic vocabulary of junior college students. *Speech Monographs*, **11**, 88-96.
- FRANCIS, W.N. (1958). *The Structure of American English* (New York).
- FRIEDERICI, A.D., SCHOENLE, P.W. and GOODGLASS, H. (1981). Mechanisms underlying writing and speech in aphasia. *Brain and Language*, **13**, 212-222.
- FRIES, C.C. (1952). *The Structure of English. An Introduction to the Construction of English Sentences* (New York).
- GIBSON, J.W., GRUNER, C.R., KIBLER, R.J. and KELLY, F.J. (1966). A quantitative examination of differences and similarities in written and spoken messages. *Speech Monographs*, **33**, 444-451.
- GOLUB, L.S. (1969). Linguistic structures in students' oral and written discourse. *Research in the Teaching of English*, **3**, 70-85.
- GOODMAN, K.S. (1972). The reading process: Theory and practice. In R.E. Hodges and E.H. Rudorf, (eds.), *Language and Learning to Read: What Teachers Should Know about Language* (Boston), pp. 143-159.
- GOODY, J. (1977). *The Domestication of the Savage Mind* (Cambridge).
- GOODY, J. (1980). Thought and writing. In E. Gellner, (ed.), *Soviet and Western Anthropology* (New York), pp. 119-133.
- GOODY, J. and WATT, I. (1963). The consequences of literacy. *Comparative Studies in Society and History*, **5**, 304-345.
- GREEN, J. (1958). A comparison of oral and written language: A quantitative analysis of the structure and vocabulary of the oral and written language of a group of college students. Ph.D. thesis, New York University.
- GREENFIELD, P.M. (1972). Oral or written language: The consequence for cognitive development in Africa, the United States and England. *Language and Speech*, **15**, 169-178.
- GRUNER, C.R., KIBLER, R.J. and GIBSON, J.W. (1967). A quantitative analysis of selected characteristics of oral and written vocabularies. *Journal of Communication*, **17**, 152-158.
- GUMPERZ, J.J. (1976). Language, communication and public negotiation. In P. Sanday, (ed.), *Anthropology and the Public Interest* (New York), pp. 273-292.

- GUMPERZ, J.J. (1977). Sociocultural knowledge in conversational inference. In M. Saville-Troike, (ed.), *Georgetown University Round Table* (Washington, D.C.), pp. 191-211.
- GUMPERZ, J.J. (1978). The conversational analysis of interethnic communication. In E.L. Ross, (ed.), *Interethnic Communication* (Auburn, Ga.), op. 13-31.
- GUMPERZ, J.J. (1980). The sociolinguistic basis of speech act theory. *Versus (Quarterni di studi Semiotici)*, **26/27**, 101-121.
- GUMPERZ, J.J. (1982). *Discourse Strategies* (Cambridge), in press.
- GUMPERZ, J.J., KALTMAN, H. and O'CONNOR, C. (1982). The transition to literacy. In D. Tannen, (ed.), *Oral and Written Discourse* (Norwood, N.J.), in press.
- GUTHRIE, J.T. (1976), (ed.). *Aspects of Reading Acquisition* (Baltimore).
- HALLIDAY, M.A.K. (1967). *Grammar, Society and the Noun* (London).
- HALLIDAY, M.A.K. (1973). *Explorations in the Functions of Language* (London).
- HALLIDAY, M.A.K. and HASAN, R. (1976). *Cohesion in English* (London).
- HANSON, V.L. (1981). Processing of spoken words: Evidence for common coding. *Memory and Cognition*, **9**, 93-100.
- HARRELL, L.E. Jr. (1957). A comparison of oral and written language in school-age children. *Mono-graphs of the Society for Research in Child Development*, **22**, (Lafayette, Indiana).
- HAWKINS, P. (1969). Social class, the nominal group and reference. *Language and Speech*, **12**, 125-135.
- HENDERSON, D. (1970). Social class differences in form-class usage among five year old children. In W. Grandis and D. Henderson, (eds.), *Social Class, Language and Communication* (London).
- HORN, E.H. (1926). A basic writing vocabulary: 10,000 words most commonly used in writing. *University of Iowa Monographs in Education*, **4**, (first series) (Iowa City).
- HOROWITZ, M.W. and BERKOWITZ, A. (1964). Structural advantage of the mechanism of spoken expression as a factor in differences in spoken and written expression. *Perceptual and Motor Skills*, **19**, 619-625.
- HOROWITZ, M.W. and NEWMAN, J.B. (1964). Spoken and written expression: An experimental analysis. *Journal of Abnormal and Social Psychology*, **8**, 640-647.
- HUDDLESTON, R.D. (1971). *The Sentence in Written English* (New York).
- HUNT, K.W. (1965). Grammatical structures written at three grade levels. *National Council of Teachers of English Research Report*, No. 3, (Champaign, Ill.).
- JONES, D. (1943). Differences between spoken and written language. *Journal of Education* (London), **70**, 207-208.
- KAUMP, E.A. (1940). An analysis of the structural differences between oral and written language of one hundred secondary school students. Ph.D. thesis, University of Wisconsin.
- KAY, P. (1977). Language evolution and speech style. In B.G. Blount and M. Sanchez, (eds.), *Socio-cultural Dimensions of Language Change* (New York), pp. 21-33.
- KEENAN, E.O. and BENNETT, T. (1977), (eds.). *Discourse Across Time and Space. Southern California Occasional Papers in Linguistics*, **5**, (Los Angeles).
- KROLL, B. (1977). Ways communicators encode propositions in spoken and written English: A look at subordination and co-ordination. In E.O. Keenan and T. Bennett, (eds.), *Discourse Across Time and Space. Southern California Occasional Papers in Linguistics*, **5**, (Los Angeles), pp. 69-108.
- LABOV, W. (1972a). *Sociolinguistic Patterns* (Philadelphia).
- LABOV, W. (1972b). *Language in the Inner City* (Philadelphia).
- LAKOFF, R.T. (1981). Some of my favorite writers are literate: The mingling of oral and literate strategies in written communication. Paper presented at the Georgetown University Round Table on Languages and Linguistics.
- LAWTON, D. (1963). Social class differences in language: A study of some samples of written work. *Language and Speech*, **6**, 120-143.
- LEVIN, H., LONG, S. and SCHAFER, C.A. (1981). The formality of the Latinate lexicon in English. *Language and Speech*, **24**, 161-171.

- MARTINET, A. (1962). *A Functional View of Language* (Oxford).
- MICHAELS, S. (1981a). Children's discourse style and differential access to literacy. Paper presented at a symposium on Urban Communication and Social Inequality, 80th Annual Meeting of the American Anthropological Association, Los Angeles.
- MICHAELS, S. (1981b). "Sharing time": Children's narrative styles and differential access to literacy. *Language in Society*, **10**, 423-442.
- O'DONNELL, R.C. (1974). Syntactic differences between speech and writing. *American Speech*, **49**, 102-110.
- O'DONNELL, R.C., GRIFFIN, W.J. and NORRIS, R.C. (1967). A transformational analysis of oral and written grammatical structures in the language of children in grades three, five, and seven. *Journal of Educational Research*, **61**, 36-39.
- OCHS, E. (1979). Planned and unplanned discourse. In T. Givón, (ed.), *Discourse and Syntax (Syntax and Semantics, Vol. 12)* (New York), pp. 51-80.
- OLSON, D.R. (1977a). From utterance to text: The bias of language in speech and writing. *Harvard Educational Review*, **47**, 257-281.
- OLSON, D.R. (1977b). Oral and written language and the cognitive processes of children. *Journal of Communication*, **27**, 10-26.
- OLSON, D.R. (1980). On the language and authority of textbooks. *Journal of Communication*, **30**, 186-196.
- ONG, W. (1980). Literacy and orality in our times. *Journal of Communication*, **30**, 197-204.
- PHILIPS, S. (1972). Participant structures and communicative competence: Warm Springs children in community and classroom. In C.B. Cazden, V.P. John and D. Hymes, (eds.), *Functions of Language in the Classroom* (New York), pp. 370-394.
- PHILIPS, S. (1975). Literacy as a mode of communication on the Warm Springs Indian Reservation. In E.H. Lenneberg and E. Lenneberg, (eds.), *Foundations of Language Development: A Multidisciplinary Approach* (New York), pp. 367-382.
- POOLE, M.E. (1972). Social class differences in language predictability. *British Journal of Educational Psychology*, **42**, 127-136.
- POOLE, M.E. (1973). Social class differences in language predictability. *Australian Journal of Education*, **17**, 300-313.
- POOLE, M.E. and FIELD, T.W. (1972). Social class and code elaboration in written communication. *Language and Speech*, **15**, 1-7.
- POOLE, M.E. and FIELD, T.W. (1976). A comparison of oral and written code elaboration. *Language and Speech*, **19**, 305-312.
- PORTNOY, S. (1973). A comparison of oral and written behavior. In K. Salzinger and R.S. Feldman, (eds.), *Studies in Verbal Behavior* (New York).
- POSNER, M.I. and HANSON, V.L. (1980). Analyzing spoken and written language. In R.A. Cole, (ed.), *Perception and Production of Fluent Speech* (Hillsdale, N.J.), pp. 201-211.
- RUBENSTEIN, H., LEWIS, S. and RUBENSTEIN, M.A. (1917). Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior*, **10**, 645-657.
- RUBIN, A. (1980). A theoretical taxonomy of the differences between oral and written language. In R.J. Spiro, B.C. Bruce and W.F. Brewer, (eds.), *Theoretical Issues in Reading Comprehension: Perspectives from Cognitive Psychology, Linguistics, Artificial Intelligence, and Education* (Hillsdale, N.J.), pp. 411-437.
- SHAUGHNESSY, M. (1977). *Errors and Expectations: A Guide for the Teacher of Basic Writing* (New York).
- SIMONS, H. (1979). Black dialect, reading interference and classroom interaction. In L. Resnick and P.A. Weaver, (eds.), *In Theory and Practice of Early Reading, Vol. 3* (Hillsdale, N.J.), pp. 111-130.
- SMITH, F. (1977). Making sense of reading — and of reading instruction. *Harvard Educational Review*, **47**, 386-395.
- SMITH, F. (1978). *Understanding Reading: A Psycholinguistic Analysis of Reading and Learning to Read* (New York).

- STUBBS, M. (1980). *Language and Literacy* (London).
- SZWED, J.F. (1981). The ethnography of literacy. In M.F. Whiteman (ed.), *Writing: The Nature, Development and Teaching of Written Communication*. Vol. 1, Variation in Writing: Functional and Linguistic-Cultural Differences (Hillsdale, N.J.), pp. 13-23.
- TANNEN, D. (1980). Spoken/written language and the oral/literate continuum. *Proceedings of the Sixth Annual Meeting of the Berkeley Linguistic Society*, pp. 207-218.
- TANNEN, D. (1981). The myth of orality vs. literacy. Paper presented at the Third Delaware Symposium on Language Studies, Newark, Delaware.
- TAYLOR, W.L. (1953). Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, **30**, 415-433.
- THORNDIKE, E.L. (1932). *A Teacher's Word Book* (New York).
- VACHEK, J. (1948). Written language and printed language. *Recueil Linguistique de Bratislava*, **1**, 65-75. (Reprinted in J. Vachek, (ed.), *A Prague school Reader in Linguistics* (Bloomington, Indiana).)
- VACHEK, J. (1976). *Selected Writings in English and General Linguistics* (The Hague).
- VOELKER, C.H. (1942). The one-thousand most frequent spoken words. *The Quarterly Journal of Speech*, **28**, 189-197.
- VYGOTSKY, L.A. (1962). *Thought and Language* (Cambridge, Mass.).
- WELLS, G. (1981). Some antecedents of early educational attainment. *British Journal of Sociology of Education*, **2**, 181-200.
- WOOLBERT, C.H. (1922). Speaking and writing – A study of differences. *The Quarterly Journal of Speech Education*, **8**, 271-285.