DOCUMENT RESUME

ED 222 857 CS 006 845

AUTHOR Ferguson, Charles A.
TITLE Cognitive Effects of Literacy: Linguistic Awareness in Adult Non-Readers. Final Report.
INSTITUTION Stanford Univ., Calif. Dept. of Linguistics.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Mar 81
GRANT NIE-G-80-0040
NOTE 22p.; For related documents see CS 006 846-847.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adult Basic Education; Adults; *Bilingual Students; English; Error Analysis (Language); Language Research; *Language Skills; Language Usage; Linguistic Theory; *Literacy; *Metacognition; *Monolingualism; Reading Ability; Reading Research; Reading Skills; Spanish Speaking; *Structural Analysis (Linguistics)
IDENTIFIERS *Metalinguistic Awareness; Papandropoulou (I); Segmentation (Verbal); Sinclair (H)

ABSTRACT

A project that investigated the metalinguistic awareness of adults (what they know about language as distinct from their ability to use language) is described and summarized in this final report. Subjects were 60 monolingual English speakers and 26 bilingual (Spanish and English) speakers enrolled in adult education classes. They fell into three groups according to reading level. Awareness of segmentation (sentences, phonemes, words) was found to be significantly related to degree of literacy, not to amount of schooling or general ability, although some components of the segmental awareness measure did not correlate with literacy. Subjects' concept of "word" as shown by elicited definitions also differed significantly by reading levels, but the differences did not fit well into the developmental scales proposed for children by I. Papandropoulou and H. Sinclair. Although the small number of subjects limited the authority of the findings, there appeared to be no significant difference in overall segmental awareness between monolinguals and bilinguals matched for reading level, age, and sex. (JL)

Cognitive Effects of Literacy:

Linguistic Awareness in Adult Non-readers

NIE Grant # NIE-G-80-0040

National Institute of Education

Principal Investigator: Charles A. Ferguson

FINAL REPORT

March 1981

Department of Linguistics
Stanford University
Stanford, CA 94305

Abstract

This project investigated certain linguistic and cognitive skills which non-literate adults bring to the task of learning to read and write, focusing specifically on their metalinguistic knowledge, i.e. their accessible knowledge about language. Sixty monolingual (English) and 26 bilingual (Spanish and English) subjects were interviewed, all of whom were students at adult learning centers in San Francisco. Average age was 28.0 and mean school grade completed 10.5. They fell into 3 groups according to reading level. Awareness of segmentation (sentences phonemes, words) was found to be highly significantly related to degree of literacy ($p < .001$) not to amount of schooling or general ability, although some components of the segmental awareness measure did not correlate with literacy. Subjects' concept of word as shown by elicited definitions also differed significantly by reading levels, but the differences did not fit well with developmental scales proposed for children (Papandropoulou & Sinclair). There appeared to be no significant difference in overall segmental awareness between monolinguals and bilinguals matched for reading level, age, and sex.

Background

The aim of this study was to investigate the linguistic and cognitive skills which non-literate adults bring to the task of learning to read and write, focusing specifically on their metalinguistic awareness, i.e. what they know about language, as distinct from their ability to use language. Many factors are involved in the acquisition of reading skills, but it seems likely that the extent of the learner's metalinguistic awareness is an important one. In

Overview

This report summarizes activities carried out during the period April 1 - December 31, 1980. The first quarter of the grant period was spent mostly in preparation: organizing the project; preparing and piloting experimental materials; designing a structured interview; training research assistants; and establishing transcription and coding systems. The second quarter was spent mainly in data collection: locating subjects, carrying out interviews, transcribing and coding interview scripts, entering into the computer those data coded for quantitative analysis. The third, final quarter was spent in completing the analyses and writing papers reporting on two aspects of the project (Barton & Hamilton 1980 and Hamilton & Barton 1980).

Day to day responsibility for the project was in the hands of the two Research Associates, Dr. David Barton and Dr. Mary E. Hamilton. Two part-time Research Associates helped with data collection and coding: William Arenio and Rebecca Horey. The project drew on the services of staff and students at six adult education centers in San Francisco, especially the San Francisco Adult Learning Center, and also made use of the facilities of the Child Phonology Project at Stanford University.

recent years a considerable and growing amount of research has been devoted to the metalinguistic awareness and related skills which preschool children can bring to the task of learning to read. This research is summarized in the volume The Child's Conception of Language (A. Sinclair et al. 1978) but there is no counterpart volume on the adult's conception of language and in fact very little research is reported on the topic.

Of the many aspects of metalinguistic awareness which can be investigated we chose to look at segmentation i.e. the recognition of such units as sentences, words, syllable, and sounds in the structure of language. Two levels of segmentation seem of particular importance for the reading process, words and phonetic segments. As noted in Read 1978, linguists have had considerable difficulty in arriving at satisfactory definitions of both these units, yet they are indicated in many of the world's writing systems including that of English and must be crucially involved in the acquisition of reading. Both have been investigated in children. In this study of non-literates we chose to deal primarily with the word, although some of the data collected deal with phonetic segments and other units.[1]

The concept of "word" is widely and frequently used by both linguists and non-linguists, but it is not a neatly defined notion. Some idea of the range of meaning of the term word in non-linguistic usage can be gotten from the list of senses given for it in any large dictionary. Two fairly concrete and probably basic senses of word for the lay person are something like 'the unit appropriate for entry in a dictionary' and 'the unit in the flow of speech or written material which is typically set off by spaces.' The first sense underlies such questions as whether house and houses are two forms of the same word or two different words or whether house, the noun, is the same word

as house, the verb. The second sense underlies such questions as whether ice cream is two words or one or whether won't is two words or one. It is the second sense that we were trying to tap in this study, since it seems more directly relevant in the acquisition of literacy.

Linguists have repeatedly tried to define the concept of word with technical precision either for describing languages or for constructing theories of language. The technical linguistic sense of word that corresponds to the second of the two non-technical senses is focused on the language unit of which sentences are comprised. (Note here that linguists have typically been concerned with possible spoken sentences in a language, not written forms of language as such.) Although they have not reached agreement on a general definition of word in this sense, they agree on the kinds of criteria to be used in the definitions: the extent to which an item can be said by itself (e.g. the -es of houses cannot normally be said by itself nor can the cran- of cranberry), the degree of substitutability of elements on either side (similar in effect to the preceding criterion), and various phonological characteristics (e.g. one primary stress per word). Linguists face two difficulties in reaching a general definition. One is that in any given language the criteria may not match, thus a highly separable item may have typical word stress or an item word-like in other respects may be limited in occurrence to a single collocation. The other difficulty is that a definition that works well for one language may not work at all for other languages, and the units identified as words differ greatly in their internal structure from one language to another. Some languages have complex derivational and inflectional morphology of suffixes and prefixes, others have little or none; some languages have extensive word compounding in which two or more words can be joined to constitute new words, others have little or none; in some languages the word is nearer to the morpheme in complexity

and separability, in other languages it is nearer to the sentence. Recent discussions of the problems involved in defining the word include Matthews 1974, Juilland 1977, Greenberg et al. 1978.

In the many writing systems that have come into use to represent language, the word and the sentence are the two units most often indicated, apart from the more purely phonological units of syllable and segmetn. In those writing systems which mark word boundaries this is done either by a word-divider symbol (as in the Old Persian cuneiform or modern Amharic orthography) or by leaving space. In the last several centuries the use of empty space as word boundary has become increasingly widespread, with the result that almost all present day orthographies around the world make at least some use of this device. In many orthographies there remain disagreements and fluctuations in the marking of word boundaries -- the two-words-or-one question -- but the general use of word space has tended to make this the primary defining characteristic of the word in the metalinguistic awareness of literate populations. Although little experimental evidence is available, it seems likely that for most readers and writers of English and many other languages judgments of word status or word boundary are made primarily in terms of whether word space is regularly used or seems appropriate. If this is so, part of the task of learning to read in any language is becoming familiar with the placing of word spaces and the notion of which strings of symbols can stand alone as words, i.e. appear with spaces on either side.

To a considerable extent orthographic words will match "linguistic" words. If the words as defined by the linguist reflect, more or less directly, speaker/hearer psycholinguistic processing units of some kind, at some level, then the task of the hearer is to match the orthographic words with these units. This

straightforward conceptualization of the process suffers from at least five uncertainities and complications. First, the linguist's identifications are not certain: as noted above, there are indeterminacies. Second, it is not clear in what way and to what extent the linguistic words are psycholinguistic processing units. Third, the degree of metalinguistic awareness of linguistic words is likely to vary greatly among individuals (the object of investigation of the present study). Fourth, orthographies typically contain a significant component of arbitrary conventions. Fifth, it may be that full recognition and awareness of words typically depends on experience with reading.

The theoretical framework with which this study proceeds includes the assumption that a gradual growth in metalinguistic awareness is a part of the overall language development of the human being. This assumption is stated in a more limited context in Ferguson & Farwell(1975), "phonological development includes the gradual development of phonological awareness: i.e. the child's ability to deal explicitly with phonological elements and relations is seen as a kind of self-discovery of his phonological organization" (438). Major purposes of research on metalinguistic awareness are the charting of this development and the discovery of factors ehich affect it. Since factors to be considered include age, general cognitive development, literacy, and knowledge of more than one language, investigation of the metalinguistic awareness of adult, non-literate subjects, monolingual and bilingual, is an appropriate area of research.

In this study we examined various aspects of the metalinguistic awareness of 60 monolingual and 26 bilingual subjects enrolled in adult education classes, by means of structured interviews. Here we report on some measures of segmental awareness and on definitions of 'word' elicited from the subjects. These results are presented in fuller form in Barton & Hamilton 1980 and Hamilton & Barton 1980

respectively. A brief note on language attitudes of the bilingual subjects is also included here as a matter of interest although it is not directly related to the metalinguistic awareness data.

Subjects and Procedures

The 86 subjects were enrolled in adult education courses at community centers in San Francisco; they were all interviewed within a month of the beginning of the courses. None of the subjects were completely non-literate: all had received some schooling. The 60 monolingual subjects fell into three groups of 20 on the basis of reading level scores: a basic group of below fourth-grade reading level, which served as the non-literate group; a higher group of above seventh-grade reading level, which served as the control literate group; and an intermediate group.

Group	# of Subjects	Reading level Mean	Range
Basic	20	2.7	1.0 - 3.9
Intermediate	20	5.8	4.0 - 7.3
High	20	9.0	7.5 - 12.0

Overall the average age was 28.0 and the mean school grade completed was 10.5 The groups did not differ significantly in age or grade completed. The 26 bilingual subjects were the total number that could be found within the time limit and the selection criteria (e.g. those with schooling only in Mexico were excluded). The bilingual subjects were then matched individually with 26 of the monolingual subjects -- primarily in terms of reading level and where possible also in terms of age and sex; the two groups did not differ significantly

in age, grade completed, or vocabulary scores.

Group	# of Subjects	Reading Level Mean	Reading Level Range
Bilinguals	26	6.9	2.4 - 12.0
Monolingual Mothers	26	6.7	2.5 - 12.0

The subjects were interviewed individually. The interviews lasted about 40 minutes and were tape recorded. Each interview consisted of questions on the recognition of sentences, phrases, and words, an oral vocabulary test; asking for a definition of words; questions asking for judgments of oral and graphic items; and background questions about schooling, language experience, and attitudes to language and education. The exact form of the interview had evolved from extensive piloting.

In the whole interview there were 62 instances where subjects could exhibit difficulty under segmentation or make errors in segmentation. The difficulties and errors shown by each subject were scored, and these items were combined in an overall measure of Segmental Awareness.

Responses to the question "What does word mean?" were put on index cards and were classified and analyzed. The cards were identified only by number which gave no clue to the reading level of the subject, so that the analysis could be done blind. A coding scheme of 25 features was devised, based on the context of the definitions and this served as the basis for the analysis.

Responses to two language attitude questions asked of the bilinguals were codified: "Do you think it's easier or more difficult for bilinguals to learn to read and write?" and "Which language do you prefer to use?" The responses were compared to the language experience of the subjects.

Results

A. Segmental Awareness

The Segmental Awareness scores varied with reading level scores, as follows:

Group	Segmental Awareness Score
Basic	13.40
Intermediate	10.35
High	7.40

The difference between the groups is significant at the $p < .001$ level. The overall Segmental Awareness score was composed of twelve component variables; five of these gave significant differences between the groups:

Component Variables	Percentage of Errors				Significance
	Basic	Intermediate	High	All	
(1) No. of Words	52.5	41.3	17.5	37.1	$p < .001$
(2) Word Identification	32.0	12.0	16.0	20.0	$p < .001$
(3) Segmentation Errors	22.0	19.3	12.0	17.8	$p < .005$
(4) Syllabification	30.0	18.0	8.0	18.7	$p < .025$
(5) Forgetting part of a sentence	10.0	7.5	1.7	6.4	$p < .02$

Of these, four (i.e. all but Syllabification) were measures of metalinguistic awareness of words. Variable (1) refers to four questions where subjects were asked for the number of words in a word or phrase (rock and roll, more or less, enough, around) and if incorrect they were asked to give the words. Subjects in the Basic group were incorrect about half the time. Variable (2) refers to six sentences where subjects were asked for specific words, such as the first, second, or last word of the sentence. No subjects in the High group made more than one error. Variables (3) and (5) refer to the same six sentences where Subjects were asked to say the sentence one word at a time.

The segmentation errors scored under component (3) were of four types: the subject splits one component word into two (e.g. always, today), the subject doesn't count the connective in a phrase (e.g. rock and roll), the subject treats two words as one (e.g. a lot, to be), the subject splits a multisyllabic word (e.g. family). Of the 250 segmentation errors, 80% were of the first two types. In these types there were significant differences between the three groups of subject, with the Basic group making the most errors (type one $p < .001$, type two $p < 2.05$).

Other segmentation phenomena in the interview material were also examined. For example the contractions didn't, he's, and everything's appeared in the test sentences, always pronounced in the contracted form by the experimenter, and the subjects differed in the ways they repeated the sentences and especially in the ways they treated the contractions when giving the sentences word for word. They could treat the contraction as a single word or as two words or, in the case of 's, they could omit the contracted part altogether (he's, he is, he). Two trends were apparent in the data for he's and everything's: more High subjects retained the contraction as a single word on both occasions and more Basic subjects omitted the 's either on both occasions or on the word-by-word question. The treatment of he's was as follows:

Group	Both 's	Both is	Both ∅	's → is	's → ∅	Other
Basic	2	2	2	6	5	3
Intermediate	8	1	0	5	4	2
High	12	0	0	6	1	1

The bilingual group was compared with the matched monolinguals on the overall Segmental Awareness measure. The bilingual's scores like those of the monolinguals, correlate significantly with reading level group, and there appears to be no difference between monolinguals and bilinguals in metalinguistic awareness as measured by this composite score.

Group	Segmental Awareness (mean no. of errors)	Correlation
Bilinguals	7.2	.57 ($p < .001$)
Monolinguals	8.3	.49 ($p < .006$)

B. Word Definitions

All but the 25 classifying features of the subjects/word definitions fall into three main types: Those referring to the functions of words, those referring to the fact that words have meaning; and those referring to the notion of units - that words consist of smaller units and/or are themselves units in larger units or system.

In spite of the small numbers involved because of the wide variety of responses, two trends can be discerned in the data when the classifications are compared with reading level groups. First, contrary to what might have been expected, subjects in the High group are more likely to refer to spoken functions while those in the Basic group are more likely to refer to writing. This difference is significant ($p < .05$ using Fisher's exact test). The numbers of subjects who included functions in their definitions are as follows:

Group	Spoken function only	Written function only	Both	Neither
Basic	4	3	3	1
Intermediate	8	2	2	2
High	11	0	1	2

Second, subjects in the High group are more likely to mention that words are part of the larger system of language, including sentences. Subjects in the High group mentioned this as opposed by those in the Basic group (Chi square = 7.2, $p < .01$). There were other non-significant trends which when combined with these results all point in the same direction that people of higher reading level tend to use a more technical vocabulary to express more linguistically sophisticated ideas.

In addition to the content feature analysis, definitions were also rated on the form and language of the definition, using a modified number of the levels analys's of Litowitz (1977), and a further analysis was done using the levels analysis of Papandropoulou and Sinclair (1974) modified to include content features not mentioned by them. The Litowitz classification goes from Level[1] "semantically empty" to Level 5 "pure Aristotelian definition". The levels are somewhat difficult to interpret in terms of the data but the scoring is as follows, with our revised labels for the levels:

Level of definition (Litowitz, adapted	Group			Totals
	Basic	Intermediate	High	
No content or unclassifiable	5	3	2	10
Listing or attitudes	5	3	2	10
Functionals	4	2	2	8
Approximations to ideas	6	11	8	25
Ideal	0	1	6	7
Totals	20	20	20	60

The largest number of responses for each group were at the Level "Approximation to ideal", and actually a majority of all the subjects were either at this levelor "Ideal". Significantly more High subjects, however, were at these levels than the other groups (chi square 26.4, $p < .05$).

Papandropoulou and Sinclair's proposed levels of development in children's definitions of words go from a Level 1 characterized by a lack of differentiation between words and referents to Level 4 in which words are autonomous units with meanings, and they hypothesize an adult Level 5 when there is formal expression of the idea that words are integrated into a system of relationships between signifiers. The scoring of the subjects,definitions by their levels is as follows:

Level of definition (Papandropoulou & Sinclair)	Group Basic	Intermediate	High	Totals
None	1	1	0	2
Level 1	0	0	1	1
Level 2	2	6	0	8
Level 3	5	1	3	9
Level 4	11	10	11	32
Level 5	1	2	5	8
Totals	20	20	20	60

This system does not differentiate clearly between any of the three groups. The most frequent kind of definition in all groups was Level 4. These results suggest that the Papandropoulou and Sinclair scale is not purely developmental, that the hypothesized Level 5 may not be typical of adults, and that other factors than degree of literacy may be important in the ability to define the word.

As part of the interview data from the bilingual subjects[2] we have their judgements on whether it is more difficult or easier for a bilingual to learn to read and write than for a monolingual, and we have their stated preference for using Spanish, English, or both. The question about difficulty of learning to read and write is so vague that the responses cannot bear much interpretation, but it is of interest to note that there is no significant variation in terms of reading level. Ten subjects responded that it was more difficult for the bilingual to learn to read and write, 10 that it was easier, 6 that it made no difference, 1 that it depended on the individual, and from one there was no response. Those who found it difficult had a median reading level of 6.9, those who found it easy had a median of 5.5, and the no difference subjects 6.25.

Of somewhat greater interest is the relation between the attitude responses and the language experience of the subjects. The bilingual subjects fell into three categories of bilingual experience: Those who had been living in the U.S. and raised bilingually from birth, those for whom Spanish was the first language (whether born in the U.A. or elsewhere) and English was acquired at school, and those of Spanish first language who had had some education in Spanish before coming to the U.S. The data may be summarized as follows:

Bilingual Experience	Judgments of difficulty for bilingual to learn to read and write				Totals	Preference for language use		
	diff	easy	no diff	other		Spanish	English	both
Bilingual from birth	4	2	3	1	10	1	9	-
Spanish first, English at school	5	5	1	1	12	4	4	4
Spanish first early education in Spanish	1	3	2	0	6	2	2	2
TOTALS	10	10	6	2	28	7	15	6

The numbers are too small and the questions too vague to draw firm conclusions from these results, but it is of interest that the U.S.-born bilinguals from birth predominately prefer English and tend to regard learning to read and write as more difficult for a bilingual, whereas the other subjects show no clear trend in language preference and those with some early education in Spanish tend to regard the bilingual's task as easier.

Conclusions

This small research study was frustrating because it is next to impossible to find completely non-literate adults in modern highly literate societies such as the U.S.A., and some hypotheses on the role of literacy in human language processing are correspondingly difficult to test in a pure form. We expected these difficulties but the frustration remained. Unexpectedly, however, it also tended to be difficult to find as many non-literate Spanish-English bilinguals as we needed to match our momolingual English subjects. In spite of these frustrations the study produced some highly suggestive results:

(1) On the basis of data from student interviews of 60 adult subjects attending adult education classes, metalinguistic awareness at the level of word segmentation i.e. recognition of words as units and the location of word boundaries, correlates well with reading level as measured by a standard battery of tests, not with age or amount of schooling. Thus this aspect of metalinguistic awareness seems not to be simply maturational. The problem remains of the directionality of the relationship: does increased literacy lead to greater awareness or is greater awareness a prerequisite to progress in literacy? This same problem of the directionality of the relationship between metalinguistic awareness and reading skills appeared in our research

on preschool children's awareness of the constituents of consonant clusters (cf. Barton et al. 1980).

(2) The developmental scales of the ability to define the word proposed by Litowitz and Papandropoulou & Sinclair are inadequate for analyzing adult definitions. As in other means of child language research, presumed adult target behavior turns out, on investigation, quite different from expectations. Ability to define words as well as the form and content of definitions elicited are complex phenomena, only in small part related to reading level.

(3) On the basis of a small number of subjects (26 bilinguals and 26 monolinguals matched for reading level, age, and sex) there is no significant difference between adult non-literate bilinguals and monolinguals in metalinguiatic awareness at the level of word segmentation, as measured by our composite Segmental Awareness scores. We had hypothesized that the bilinguals, having experience in word segmentation in two languages would be significantly superior to monolinguals; other researchers who look for cognitive disadvantages for bilinguals might have hypothesized the converse. Neither hypothesis was confirmed. This issue is of sufficient importance that further research on larger numbers of subjects under varied conditions of bilingualism should be pursued.

(4) On the basis of self-report data from 28 Spanish-English bilinguals of three categories (10 bilingual from birth; 12 Spanish first, English acquired at school; 6 Spanish first and early schooling in Spanish-speaking country) the bilinguals from birth prefer English rather than Spanish or "both" (9 to 1) and tend to regard learning to read and write as more difficult for a bilingual than for a monolingual. The other subjects show no language preference and the subjects with early schooling in Spanish tend to regard the bilingual's task as easier. This finding suggests, although very tentatively considering

the small number of subjects and the self-report nature of the data, that American Spanish-English bilinguals from birth have a self-image of being disadvantaged and that early instruction in the mother tongue may facilitate literacy development.

Other data from the interviews unanalyzed at the end of the grant period will be made available for study by graduate students at Stanford interested in metalinguistic awareness, literacy, bilinguals, or bilingual education.

Footnotes

[1]This study was regarded as part of a series of studies undertaken under the auspices of the Child Phonology Project relating to different aspects of metalinguistic awareness. For a report on another study in the series, cf. Barton et al. 1980. A more general theoretical framework was presented in Ferguson 1980. The Child Phonology Project at Stanford is a continuing operation, focusing on various aspects of phonological development in children; it has been supported by grants from the National Science Foundation, the National Institute of Health, the National Institute of Education, the William T. Grant Foundation, and Stanford University.

[2]Attitudinal and other data are available from 28 bilingual subjects, i.e. two in addition to the 26 matched with monolinguals.

References

Barton, D., Miller, R. & Macken, M.A. 1980. Do children treat clusters as one unit or two? Papers and Reports on Child Language Development 18.105-137.

Ferguson, C.A. & Farwell, C.B. 1975. Words and sounds in early language acquisition. Language 51.419-439.

Greenberg, J. H., Ferguson, C.A. & Moravscik, E. (eds.) 1978. Universals of human language, Vol. 3 Word Structure. Stanford, CA: Stanford University Press.

Juilland, A.G. 1977. The decline of the word. Saratoga, CA: Anma Libri.

Litowitz, B. 1977. Learning to make definitions. Journal of Child Language 4.289-304.

Matthews, P.H. 1974. Morphology: an introduction to the theory of word structure. London: Cambridge University Press.

Papandropoulou, I. & Sinclair, H. 1974. What is a word? Human Development 17.241-258.

Sinclair, A., Jarvella, R.J. & Levelt, W.J.M. 1978. The child's conception of language. New York: Springer Verlag.

Papers and Oral Presentations Resulting from the Study

The following papers and oral presentations resulted directly from the research of the project:

1. Barton, D. & Hamilton, M.E. 1980. Awareness of the segmental structure of English in adults of various literacy levels.

2. Ferguson, C.A. 1981. Language awareness and literacy. Lecture given at the School of Audiology and Speech Sciences, University of British Columbia, Vancouver B.C., Canada, February 12, 1981.

3. Hamilton, M.E. & Barton, D. 1980. A word is a word: metalinguistic skills in adults of varying literacy levels. To appear in a slightly revised version in the Journal of Psycholinguistic Research.

4. Hovey, R. 1980. Bilingual adults' attitudes to learning to read and write. Term paper based on the project.

Copies of items 1. and 3. are included with this report.