

WORKING PAPERS IN LINGUISTICS No. 22

PAPERS IN PSYCHOLINGUISTICS

AND SOCIOLINGUISTICS

Edited by Olga K. Garnica

Papers By

Mary Louise Edwards, Sara Garnes,

Olga K. Garnica, Roy C. Major,

and John Perkins

Department of Linguistics

The Ohio State University

1841 Millikin Road

Columbus, Ohio 43210

February, 1977

OSU Working Papers in Linguistics No. 22

Papers in Psycholinguistics
and Sociolinguistics

The papers in this volume examine some important issues in psycholinguistic and sociolinguistic research--language development, bilingualism, linguistic variation. The paper by Major is a revision of his OSU M.A. thesis (1976). My contribution on the prosodic aspects of speech to young children is a slightly revised version of my Stanford Ph.D. dissertation (1975).

OKG

Table of Contents

	Page
List of Working Papers in Linguistics	iii
Sara Garnes, "Effects of Bilingualism on Perception"	1
Olga K. Garnica, <u>Some Prosodic Characteristics of Speech to Young Children</u>	11
Olga K. Garnica, "Nonverbal Concomitants of Language to Children: Clues to Meaning"	73
Olga K. Garnica and Mary Louise Edwards, "Phonological Variation in Children's Speech: The Trade-Off Phenomena"	81
Roy C. Major, <u>Phonological Differentiation of a Bilingual Child</u>	88
John Perkins, "A Sociolinguistic Glance at the Great Vowel Shift of English"	123

List of Working Papers in Linguistics

- No. 1, December 1967 out of print
 Articles by Dale Elliott, Charles Fillmore, James Heringer, Terence Langendoen, Gregory Lee, Ilse Lehiste, and Sandra Thompson.
- No. 2, November 1968 (OSU-CISRC-TR-68-3) PB-182 596
 Articles by Charles Fillmore and Ilse Lehiste.
- No. 3, June 1969 (OSU-CISRC-TR-69-4) PB-185 855
 Articles by Dale Elliott, Shuan-fan Huang, Terence Langendoen, Gregory Lee, and Ilse Lehiste.
- No. 4, May 1970 (OSU CISRC-TR-70-26) PB-192 163
 Articles by Gaberell Drachman, Mary Edwards, Charles Fillmore, Gregory Lee, Patricia Lee, Ilse Lehiste, and Arnold Zwicky.
- No. 5, June 1969 out of print
Twana Phonology, by Gaberell Drachman.
- No. 6, September 1970 (OSU-CISRC-TR-70-12) PB-194 829
 Articles by Charles Fillmore, Ilse Lehiste, David Meltzer, Sandra Thompson, and Marcel Tatham.
- No. 7, February 1971 (OSU-CISRC-TR-71-7) PB-198 278
 Articles by Alexander Grosu and Gregory Lee.
- No. 8, June 1971 (OSU-CISRC-TR-71-7) PB-202 724
 Articles by Dale Elliott, Michael Geis, Alexander Grosu, Barry Nobel, Ann Zwicky, and Arnold Zwicky.
- No. 9, July 1971 (OSU-CISRC-TR-71-8) PB-204 002
 Articles by Zinny Bond, Richard Gregorski, Andrew Kerek, Ilse Lehiste, Linda Shockey, and Mary Wendell.
- No. 10, August 1971 Edited by Charles Fillmore out of print
 Articles on Case Grammar.
- No. 11, February 1972 ED 062 850
 Articles by James Heringer, Patricia Miller, Lawrence Schourup, and Richard Wojcik.
- No. 12, June 1972 (OSU-CISRC-TR-72-6) PB-210 781
 Articles by Richard Gregorski, David Meltzer, Ilse Lehiste, and Linda Shockey.

- No. 13, December 1972 ED 077 268
Alexander Grosu, The Strategic Content of Island Constraints.
- No. 14, April 1973 ED
Studies in Phonology and Methodology. Articles by James Hutcheson, Ronald Neeld, Mieko Ohso, Lawrence Schourup, Holly Semiloff, Clare Silva, and Arnold Zwicky.
- No. 15, April 1973 ED 082 566
Articles by Angeliki Malikouti-Drachman, Gaberell Drachman, Mary Edwards, Jonnie Geis, and Lawrence Schourup.
- No. 16, December 1973 ED
Mostly Syntax and Semantics. Articles by Michael Geis, Fred Goodmar, Patricia Lee, Ronald Neeld, Jerrold Sadock, Clare Silva, and Arnold Zwicky.
- No. 17, May 1974 \$2.00
Articles by Sara Garnes, Ilse Lehiste, Patricia Miller, Linda Shockey, and Arnold Zwicky.
- No. 18, June 1975 \$2.00
Articles by Michael Geis, Sheila Geoghegan, Jeanette Gundel, Ronald Neeld, Geoffrey Pullum, and Arnold M. Zwicky
- No. 19, September 1975 Edited by Robert Herbert \$2.50
Patterns in Language, Culture, and Society: Sub-Saharan Africa. Proceedings of the Symposium on African Language, Culture and Society.
- No. 20, September 1975 Edited by Robert Herbert \$2.50
Proceedings of the Sixth Conference on African Linguistics.
- No. 21, May 1976 Edited by Arnold Zwicky \$2.50
Papers on Nonphonology. Articles by Steven Boër, Marion Johnson, Robert Kantor, Patricia Lee, William Lycan, and Jay Pollack.
- No. 22, February 1977 Edited by Olga Garnica \$3.50
Papers in Psycholinguistics and Sociolinguistics. Articles by Mary Louise Edwards, Sara Garnes, Olga K. Garnica, Roy C. Major, and John Perkins.

Orders for those issues still available (indicated by price) should be sent to:

OSU WPL
Department of Linguistics
Ohio State University
1841 Millikin Road
Columbus, OH 43210

Issues with PB numbers are available through:

The National Technical Information Center
The U.S. Department of Commerce
5285 Port Royal Road
Springfield, Virginia 22151

Issues with ED numbers--11, 13, parts of 14 (articles by Holly Semiloff-Zelasko, "Glide Metatheses" [ED 104 140] and Clare M. Silva, "Metathesis of Obstruent Clusters" [ED 105 735]), 15, and parts of 16 (articles by Clare M. Silva, "Adverbial -ing" [ED 095 720], Clare M. Silva and Arnold M. Zwicky, "Discord" [ED 095 719], and Arnold M. Zwicky, "Bibliography I, Coivs" [ED 096 826] and "Bibliography III, Forestress and Afterstress, Compounds and Phrases" [ED 095 721]) are available through:

ERIC Clearinghouse on Language and Linguistics
Center for Applied Linguistics
1611 North Kent Street
Arlington, Virginia 22209.

Some Effects of Bilingualism on Perception*

Sara Garnes

1. Statement of the problem.

In this paper I present the results of an experiment designed to investigate some of the effects of bilingualism on perception. The subjects were Icelandic monolinguals and Icelandic-English bilinguals (see discussion in the next section). The phonetic parameter under investigation was duration.

In Icelandic, quantity is phonemic and is distributed over the syllable: a long vowel is followed by a short consonant or a short vowel is followed by a long consonant, e.g. ís [i:s] 'ice' vs. íss [is:] 'of ice'; vaka [va:ka] 'to wake' vs. vagga [vak:a] 'to rock' (see Garnes, in press, for discussion). In contrast, English has no phonemic quantity opposition. Vowels before voiceless stops are two-thirds the duration of vowels before voiced stops; however, this difference is sub-phonemic, e.g. bet [bɛt] vs. bed [bɛ:d] (Peterson and Lehiste 1960).

Since the second language, English, has a different quantity system than the native language, it could be expected that the perceptual responses of the bilinguals would fall into one of three categories predicted by three different hypotheses. The null hypothesis predicts no interference between the two languages. According to this hypothesis, each linguistic system is self-contained so that the perceptual responses of the monolinguals and of the bilinguals will not be different. The other two hypotheses both predict that contact with a second language will produce a change in perception. The second hypothesis predicts that bilinguals will perceive the contrast in their native language more accurately than do the monolinguals. Since the bilinguals have been exposed to a contrasting system their perception of the opposition in their native language will be heightened. We can call this the awareness hypothesis. The third hypothesis predicts that bilinguals will perceive the contrast in their native language less accurately than do the monolinguals. This hypothesis, which we can call the confusion hypothesis, predicts that the bilinguals' perception will be confused due to a conflict between the two different linguistic systems.

2. Method.

In order to determine which of the three hypotheses is the correct one, a listening test was administered which consisted of stimuli based on the minimal pair, ís [i:s] 'ice' with a long vowel and a short consonant, and íss [is:] 'of ice' with a short vowel and a long consonant. Thirteen vowel durations ranging from 80 to

320 ms. in 20 ms. increments were combined with 13 consonant durations ranging from 200 to 440 ms., also in 20 ms. steps. Stimuli were prepared on a Glace-Holmes terminal analog synthesizer. The total 169 stimuli were randomized and spliced together into groups of ten with 5 seconds between individual stimuli and 10 seconds between groups.¹

Two groups of subjects were tested. All were natives of Iceland and were college students at the University of Iceland or at universities in the U.S., or of that generation, at the time of testing. Since no Icelandic university students are strictly monolingual in the sense that they have never studied a foreign language, the distinction was made that to be considered monolingual, the subjects must not have studied out of Iceland. Seventeen monolinguals participated, 9 males and 8 females. The 9 bilingual subjects, 6 males and 3 females, were fluent speakers of English and had resided out of Iceland for at least one school year.² The forced-choice listening test was administered on a Uher 4000 Report-L tape recorder in quiet surroundings. Answers were recorded on a prepared answer sheet. The listening test was taken twice by both groups of subjects yielding two responses per stimulus from each subject.

3. Results.

The results of the listening test show a basic similarity between the two groups: vowel length is a more salient cue than consonant length in the perception of the quantity opposition. For stimuli with the shorter vowel durations all subjects agreed on íss [i:s] 'of ice', whereas with the longer vowel durations all subjects agreed on ís [i:s] 'ice'. Thus, for all 13 consonant lengths cross-over between the two lexical categories obtained, as is illustrated in Figure 1. The vertical dimension on the graphs shows the percentage of judgments for íss, the inverse of the judgments for ís. The vowel durations are given on the ordinate. The graphs show that vowel length is the most important cue in perceiving the quantity distinction.

However, consonant length also contributes to perception of the quantity contrast. As the consonant length increases, the vowel length required for cross-over between the lexical categories also increases. For example, in Figure 1 with a consonant length of 200 ms., cross-over between lexical categories occurs at ca. 155 ms. vowel length, but at a consonant length of 440 ms., as illustrated in the last graph in Figure 1, cross-over occurs at ca. 190 ms. vowel length. Thus it appears that a variable ratio of vowel to consonant length actually accounts for the responses for both the monolingual and bilingual subjects.

Figure 1. Responses to the listening test arranged by increasing consonant length from consonant length of 200 ms. to 440 ms. Vowel length appears on the ordinate ranging from 80 ms. to 320 ms. Interrupted lines show bilingual responses; connected lines show monolingual responses.

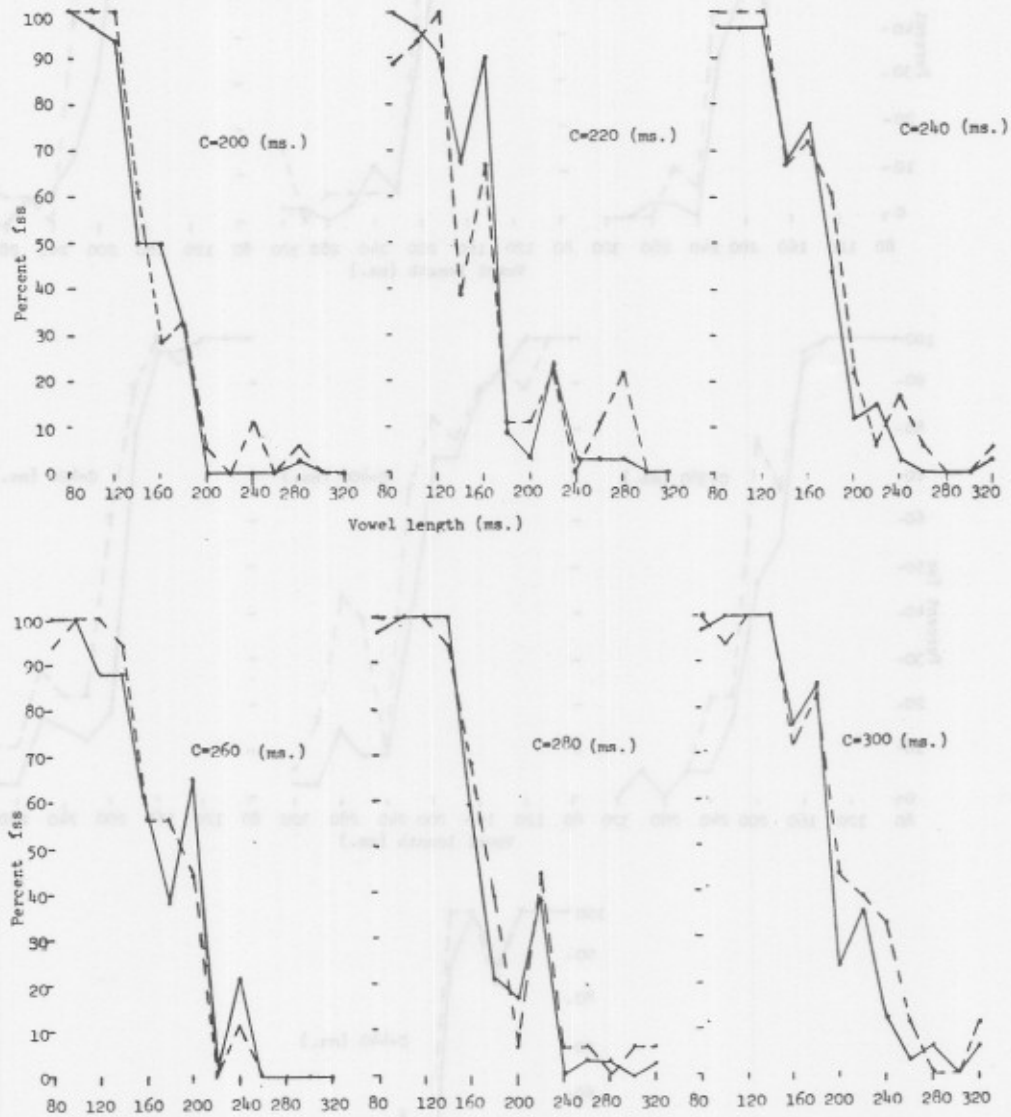
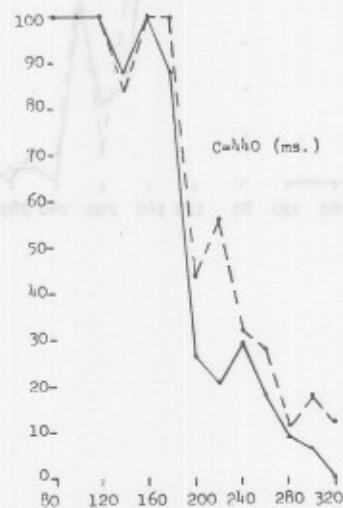
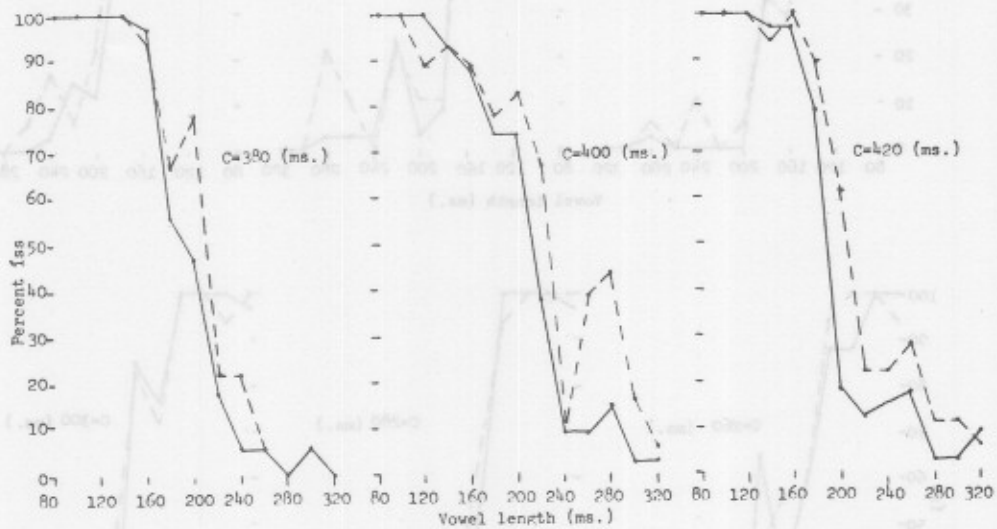
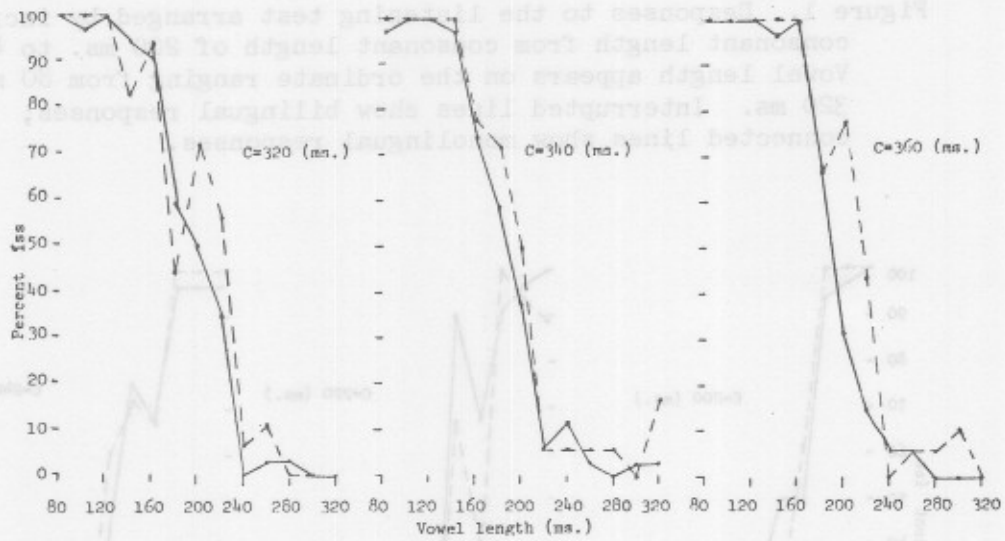


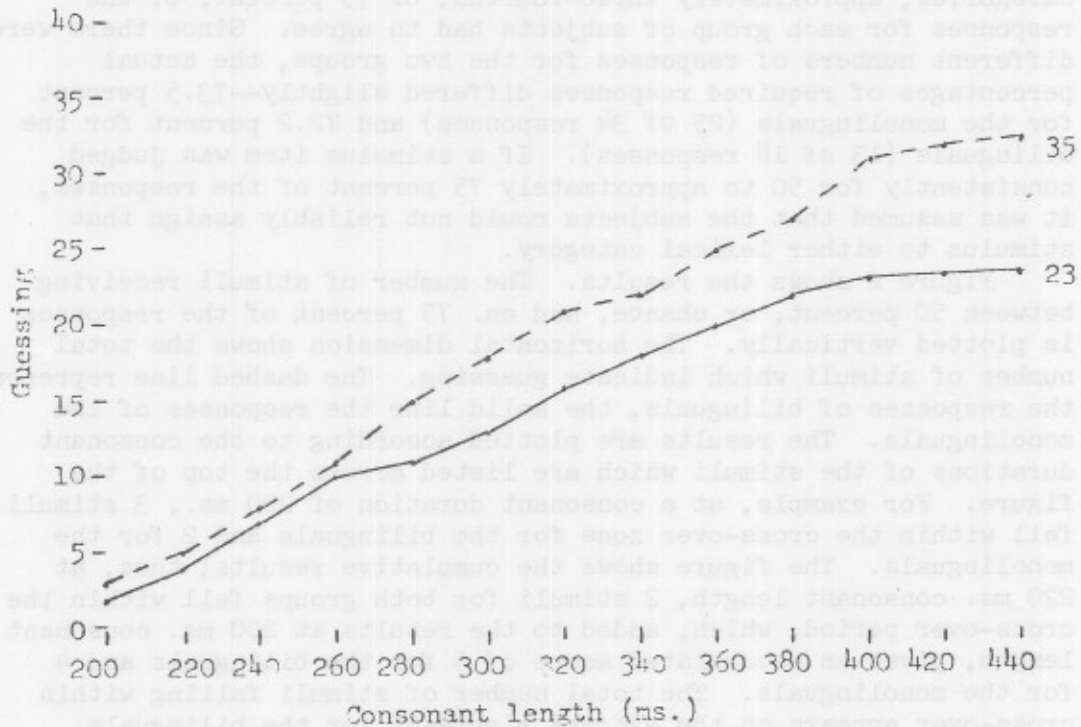
Figure 1. (continued)



The differences between the two groups of subjects are due to the different proportions of stimuli which indicate guessing between the two lexical items. The responses to the listening test were analyzed according to the following criteria. In order to consider that a stimulus was assigned to one of the two lexical categories, approximately three-fourths, or 75 percent, of the responses for each group of subjects had to agree. Since there were different numbers of responses for the two groups, the actual percentages of required responses differed slightly--73.5 percent for the monolinguals (25 of 34 responses) and 72.2 percent for the bilinguals (13 of 18 responses). If a stimulus item was judged consistently for 50 to approximately 75 percent of the responses, it was assumed that the subjects could not reliably assign that stimulus to either lexical category.

Figure 2 shows the results. The number of stimuli receiving between 50 percent, or chance, and ca. 75 percent of the responses is plotted vertically. The horizontal dimension shows the total number of stimuli which indicate guessing. The dashed line represents the responses of bilinguals, the solid line the responses of the monolinguals. The results are plotted according to the consonant durations of the stimuli which are listed across the top of the figure. For example, at a consonant duration of 200 ms., 3 stimuli fell within the cross-over zone for the bilinguals and 2 for the monolinguals. The figure shows the cumulative results; thus, at 220 ms. consonant length, 2 stimuli for both groups fell within the cross-over period, which, added to the results at 200 ms. consonant length, gives an accumulated score of 5 for the bilinguals and 4 for the monolinguals. The total number of stimuli falling within cross-over appears on the extreme right--35 for the bilinguals versus 23 for the monolinguals.

Figure 2. Number of stimuli receiving between 50% to ca. 75% of the responses.



Statistical analysis of these results shows them to be highly significant. According to the Wilcoxon matched-pairs signed-ranks test, the experimentally obtained results are significant at the .001 level. This predicts that the differences found here between the monolinguals and the bilinguals would occur only once in 1000 times if attributed simply to chance.

5. Discussion.

The results clearly do not support the null hypothesis. On the contrary, the bilinguals do respond quite differently from the monolinguals, indicating that the two linguistic systems interact and affect the perception of the bilinguals. According to the second hypothesis, the awareness hypothesis, the bilinguals would have responded more accurately than the monolinguals. This hypothesis must also be rejected since the bilinguals agreed less than the monolinguals in assigning stimuli to either lexical item. However, significant support is found for the third hypothesis which predicts confusion in the perception of the bilinguals. Apparently, mastery

of a foreign language, which has a different quantity system than the native language, affects the perceptual behaviour of the bilinguals.³

6. Implications.

There are several implications of this study for linguistic theory. Of interest for psycholinguistics is another documentation that the mental representations of two linguistic systems can overlap. In the experienced bilingual, the foreign language apparently interferes with the mental representation of the native language so that the phonetic representation of the quantity opposition differs between the monolingual and bilingual populations.

Caramazza, et al (1973), have reported the results of another phonetic study which deals with the voice onset time in bilingual individuals. Three groups of speakers were used: monolingual Canadian French, monolingual English, and bilingual Canadian French and English. The bilinguals were tested in both Canadian French and English speech modes. Voiced-voiceless judgments were obtained for stimuli which varied along the voice onset time continuum. Stops corresponding to three places of articulation--bilabial, alveolar and velar--in initial position before the vowel [a] were tested. Canadian French is characterized as having unaspirated voiceless stops, e.g. [pa], while English has aspirated initial voiceless stops, e.g. [p^ha]. The results of the experiment show a clear distinction between the duration of voice onset time required for the monolinguals to cross-over between the perception of voiceless and voiced stops in the respective languages. The perceptual cross-overs of the bilingual individuals fall between the perceptual cross-over of the monolingual subjects whether they are in the Canadian French or in the English speech mode. Caramazza, et al., also investigated the corresponding aspects of the production of the 4 subsets of speakers. The results indicate that the bilinguals are "better able to adapt their production mechanisms than their perceptual mechanisms to the second language." Thus it appears that "language switching is easier for production than for perception" (Caramazza, et al., 1973:427).

Using the same type of consonant-vowel stimuli, Williams (1974) reports on the perception and production of word-initial voiced and voiceless labial stops by monolingual English, monolingual Puerto Rican Spanish, and bilingual speakers of English and Puerto Rican Spanish. The segmental differences between Canadian French and Puerto Rican Spanish are similar in that voiceless stops are unaspirated. The perception of the bilingual English-Puerto Rican Spanish speakers in Williams' study shows confusion as opposed to the monolinguals' perception. Thus, there appears to be a fair amount of evidence indicating a difference between bilinguals' and monolinguals' perception of aspirated voiceless, unaspirated voiceless and voiced stops.

In general, the bilinguals' production has not been shown to become significantly different from the monolinguals' production. Although I have not systematically investigated the production of the quantity contrast for all the Icelandic subjects, preliminary investigations show no apparent systematic differences between the production of the quantity contrast between monolingual and bilingual speakers, other things being equal.

Kolers (1968) reports the results of a series of experiments dealing with syntactic and lexical effects of bilingualism. He found that subjects who read passages of mixed English and French words based on either English or French syntax took the same time when they silently read passages in either French or English. However, when asked to read aloud, the subjects took longer--ca. 33 cs. for each code-switch. In an experiment designed to determine the relationship between meaning and lexical storage, he found that words representing concrete objects such as lamb, thorn and tree, were readily accessible in either of the two languages but that the accessibility of words which were more abstract such as love/Liebe and democracy/Demokratie was closely bound to the language by which they were encoded. The experiment was based on the observation that the more often a word appears in a series of words, the more likely a monolingual subject is to recall the word. This same result was obtained whether the word was presented in the first or second language for bilingual individuals, e.g. English fold vs. French pli. Kolers' experiments show the complex and inter-related nature of some non-phonetic parameters in the bilingual individual.

The results of these experiments indicate that a detailed linguistic history of subjects needs to be considered in selecting subjects for participation in perception tests. Possible inter-language effects need to be considered in evaluating responses of bilingual individuals.

Implications for sociolinguistics are also relevant, especially in the field of languages in contact. The linguistic relationships among adstratum, substratum and superstratum languages have long been discussed (Weinreich 1974 [1953]). The results of these experiments show the psychological reality of some of the affects of languages in contact.

The line between languages in contact and language change is not a distinct one. Thus, the study of languages in contact is of interest to both socio- and historical linguists. The results of the experiment reported here could be interpreted as providing a mechanism for language change. What has been interpreted here as the bilingual speakers' indecision about the phonemic boundaries in their native language may very well lead to the acceptance of a broader variation in phonemic distinctions. If this is true, language change may proceed through several quite complicated steps: in step 1 bilinguals' perception of a given contrast becomes confused, as we have seen in this study. In step 2 bilinguals become more tolerant of variation. Since it is tolerated, more variation

actually occurs in step 3. In step 4 old boundaries begin to shift, in step 5 new boundaries are established as the old distinctions fade away.

This possible mechanism of language change presents a problem: how could the bilingual experience be transmitted to successive generations? Perception is a very personal matter; we are not able to verbalize about criteria used in making perceptual judgments. Although highly speculative, one possible mechanism for transmission is through language acquisition. Thus, bilingual parents may be permissive parents, linguistically speaking. Children of bilingual speakers may acquire a wider variation in producing contrasts. Their parents' boundaries would be less strict; thus the parents would be more likely to accept a wider variety of pronunciations. If this environmental feature is repeated over several generations by sufficient numbers of speakers, language change may occur.

7. Conclusion

In conclusion, the present study shows a significant difference in the perception of the quantity system of Icelandic between monolingual and bilingual native speakers of the language. Hopefully, the results will be helpful in increasing our understanding of the complex nature of the mental representation of language and languages, as well as understanding the nature of the influence of languages in contact on phonological processes in language change.

Footnotes

*This research was supported in part by NSF Grant GS-36252. I wish to thank Höskuldur Thráinsson for his assistance in helping to administer some of the listening tests, the College of Humanities for funding the computer time and the staff at the Instructional and Research Computer Center for their cooperation. An earlier version of this paper was presented at the 66th meeting of the Society for the Advancement of Scandinavian Studies held in Austin, Texas in April 1976. I appreciate comments from R. Austerlitz, Höskuldur Thráinsson, A. S. Liberman and Z. S. Bond on that version of the paper.

1. For a detailed description of the stimuli, see Garnes, in press.

2. There are less than 200,000 native speakers of Icelandic and the number of qualifying bilingual speakers I have been able to locate is quite small. In the autumn of 1976 I hope to be able to administer the listening test to 4 additional bilingual subjects in Ann Arbor, Michigan.

3. It might be suggested that the reason for the increased guessing was due to an increased sensitivity on the part of the bilinguals; i.e. it is precisely their heightened awareness which creates the confusion. Although this may be the source of the

increase in guessing, the fact remains that the bilinguals in this study responded less unanimously than the monolinguals.

References

- Caramazza, A., G. H. Yeni-Komshian, E. B. Zurif, and E. Carbone. 1973. The acquisition of a new phonological contrast: the case of stop consonants in French-English bilinguals. *JASA* 54.421-428.
- Garnes, Sara. In press. *Quantity in Icelandic: Production and Perception*. Hamburger Phonetische Beiträge. Buske Verlag. Hamburg.
- Kolers, Paul A. 1968. Bilingualism and information processing. *Scientific American* 218.78-86.
- Peterson, Gordon E. and Ilse Lehiste. 1960. Duration of syllable nuclei in English. *JASA* 32.693-703.
- Weinreich, Uriel. 1974 [1953]. *Languages in Contact*. Mouton. The Hague.
- Williams, Lee. 1974. *Speech Perception and Production as a Function of Exposure to a Second Language*. Unpublished complex dissertation. Harvard University.

SOME PROSODIC CHARACTERISTICS OF SPEECH
TO YOUNG CHILDREN

A DISSERTATION
SUBMITTED TO THE COMMITTEE ON LINGUISTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

By
Olga Kaunoff Garnica

November 1974

SOME PROSODIC CHARACTERISTICS OF SPEECH
TO YOUNG CHILDREN

A DISSERTATION
SUBMITTED TO THE COMMITTEE ON LINGUISTICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
c 1974

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

By
Olga Kaunoff Garnica

DOCTOR OF PHILOSOPHY

By

Olga Kaunoff Garnica

November 1974

Preface

I wish to gratefully acknowledge the assistance of the following persons whose advice and support at various points in the progress of this work helped to bring it to a fruitful completion:

--Eve V. Clark, Charles A. Ferguson, and Ilse Lehiste, my principal advisors, for their expertise and encouragement;

--Dorothy Huntington and Clara Bush, who were extremely helpful members of my oral examination committee;

--John Ohala, director of the Phonetics Laboratory at the University of California, Berkeley for the use of the facilities and consultant personnel;

--the director of the language laboratory at Stanford University for the use of tape recording facilities and indispensable technical assistance;

--the mothers and children who participated in the study;

--many of my fellow graduate students at Stanford University who were available to exchange ideas and provide encouragement.

--finally, to my family and very especially to my husband for his immeasurable help and encouragement especially thru some of the more trying periods.

TABLE OF CONTENTS

	Page
Preface	13
List of Tables and Figures	16
 CHAPTER	
1. INTRODUCTION	
1.1. Introduction	17
1.2. Speech registers	18
1.2.1. Definition of concept	18
1.2.2. "Simplified" registers	20
1.3. Baby talk register	23
1.3.1. Types of studies	23
1.3.2. Phonological and lexical features	24
1.3.3. Syntactic complexity and redundancy features	25
1.3.4. Prosodic features	27
1.4. Hypotheses	29
 2. METHOD	
2.1. Subjects	31
2.2. Procedure	31
2.2.1. Testing sessions	31
2.2.2. Verbal tasks	32
2.2.2.1. Picture task	32
2.2.2.2. Puzzle task	34
2.2.2.3. Story reading task	35
2.2.3. Interviews	36
2.2.4. Data collection	37
2.2.4.1. Taping of sessions	37
2.2.4.2. Physical setting	37
2.2.4.3. Instrumentation	38
2.2.5. Data processing	38
2.2.5.1. Selection of utterances	38
2.2.5.2. Acoustic analysis	39
2.2.5.3. Perceptual analysis	40
2.2.5.4. Measurement of fundamental frequency and duration	40
 3. RESULTS	
3.1. Average fundamental frequency	42
3.2. Frequency range	45
3.3. Falling vs. rising sentence final pitch terminals	49
3.4. Use of whispering	50

	Page
3.5. Duration	51
3.6. Distribution of primary stress	54
3.7. Summary of results	55
4. DISCUSSION	
4.1. Functions of characteristics of the baby- talk register	56
4.1.1. Prosodic features: Social function . .	57
4.1.2. Prosodic features: Analytic function .	58
4.2. Age of child listener	60
4.3. Speaker interviews	61
4.4. Implications for further research	64
5. CONCLUSION	
5.1. Conclusion	66
5.2. Implications for theory of language development	66
5.3. Implications for linguistic theory	67
REFERENCES	69

FIGURE

1. Mean fundamental frequency level for G2 subjects by type of listener	43
2. Mean fundamental frequency level for G1 subjects by type of listener	44
3. Mean fundamental frequency level by subject group and type of listener	45
4. Comparison of frequency ranges (in sentences) for G1 subjects by type of listener	47
5. Comparison of frequency ranges (in sentences) for G2 subjects by type of listener	47
6. Frequency range for G1 subjects by type of listener . .	48
7. Frequency range for G2 subjects by type of listener . .	49

List of Tables and Figures

TABLE		Page
1.	Age and sex of children participating with their mothers in the study	31
2.	Description of materials in picture task	33
3.	Comparison of fundamental frequency and frequency range data for C1 and C2 subjects	42
4.	Average fundamental frequency for C1 and C2 subjects by session	42
5.	Frequency range for C1 subjects by type of listener	46
6.	Frequency range for C2 subjects by type of listener	46
7.	Number of sentences with rising pitch terminals in the child-listener session by subject group and task	50
8.	Average duration (in sec) of verbs in Puzzle Task	52
9.	Average duration (in sec) of color terms in Puzzle Task	52
10.	Number of instances of two primary stresses per sentence in six sentences from the Puzzle Task by type of listener	54
11.	Presence and absence of some prosodic characteristics in adult speech by function and age of child listener	60

FIGURE

1.	Mean fundamental frequency level for C2 subjects by type of listener	43
2.	Mean fundamental frequency level for C1 subjects by type of listener	44
3.	Mean fundamental frequency level by subject group and type of listener	45
4.	Comparison of frequency ranges (in semitones) for C1 subjects by type of listener	47
5.	Comparison of frequency ranges (in semitones) for C2 subjects by type of listener	47
6.	Frequency range for C1 subjects by type of listener	48
7.	Frequency range for C2 subjects by type of listener	49

CHAPTER 1

1.1. Introduction

It is now often claimed that speech addressed to children and speech addressed to adults differ in systematic and identifiable ways. This view represents a reversal of the direction taken by Chomsky (1967) and others (McNeill, 1970) who held the view that adult speech was mostly ungrammatical, replete with false starts, hesitations and slips of the tongue regardless of the addressee. This was supported to a certain degree empirically by Bever, Fodor, and Weksel (1965) who made calculations of grammatical and ungrammatical sentences in several speech samples.

This position had a strong influence on the theory of language development that was proposed. If the verbal input to the child is fragmented, confusing, and to a great degree unsystematic, then acquisition probably occurs independently of the linguistic environment. The input to the child must play a minimal role in the acquisition of language. Some verbal input was necessary, of course. This was clear from cases of children totally deprived of the opportunity to hear language. But input was relegated to a secondary role. Much of language acquisition was attributed to the child's innate capacities. The reasoning was that, if the language the child hears or overhears is not of a form that reveals the underlying systematicity of language, then some other mechanism must be available to the child in order for him to be able to abstract rules and gain knowledge about this systematicity.

Not all persons held this view. One of the strongest opponents was Labov (1970). He pointed out:

The ungrammaticality of everyday speech appears to be a myth with no basis in actual fact. In the various grammatical studies that we have conducted, the great majority of utterances--about 75%--are well formed sentences by any criterion. When rules of ellipsis are applied and certain universal editing rules to take care of stammering and false starts, the proportion of truly ungrammatical and ill-formed sentences falls to less than two percent. (p. 42).

Although Labov makes no specific reference to speech directed to children, others had noted that the language heard by children is likewise neither phonologically nor grammatically deviant (Brown and Bellugi, 1964; Waterson, 1971).

What are the consequences of this alternative view of the role of the verbal environment for a theory of language development? One possible consequence is that greater emphasis is placed on the mother-child verbal interaction and the contribution of this experience to the child's learning of language. A direct result of this is that the importance of the innate mechanisms is deemphasized.

Another consequence is that the child's role becomes more important since the child can be seen as a more active participant in the language development process.

There are two questions relating to the verbal input to the child and his language development. First, what are the linguistic characteristics of speech addressed to the child? This is a purely descriptive question but necessarily preliminary to the second question: what features of the verbal environment are critical for learning language? Certainly, the latter question is more relevant to a theory of language development, and, more broadly, to the problem of cognitive development. Its answer, however, depends to some extent on the answer to the first question.

The present investigation seeks in part to answer the first question. It focuses on one linguistic aspect of adult speech to the young child, the prosodic and paralinguistic features of such speech, in an attempt to specify it more fully. In this sense, the study is intended to contribute at least indirectly to a more adequate theory of language development.

Finally, the findings of this study may contribute generally to a more adequate linguistic theory. Adult speech directed to children can be compared to other kinds of systematic speech adjustments adults make in response to different addressees in different social situations, i.e., speech registers (see 1.2.). Knowledge about the linguistic features characteristic of these adjustments and the distribution of their use will provide information which will have to be accounted for eventually by linguistic theory.

1.2. Speech Registers.

1.2.1. Definition of concept.

The term speech register refers to a type of social variation in language which is defined by the use to which language is put. Register is the most commonly used term, although the terms speech variety or speech style are also sometimes employed. All these are equally acceptable with perhaps the exception of the word "style" which has many strong associations with literature study and literary criticism that are not appropriate to the present use of the term. An excellent overview of the notion of speech register is available in Ellis and Ure (1969; and references therein). What follows is an elementary and highly schematic discussion of the basic points on the concept of register as preparation for the subsequent discussion.

If we assume that the focus of linguistic study is linguistic behavior, the reasonable object of study is language. In studying "language" we can be concerned with (a) different codes (tongues, languages), such as English, Chinese, etc., or (b) regional varieties within a single code, such as Boston English, Brooklyn English, Norfolk English, etc., or (c) social class varieties of a particular regional variant, such as the English of lower/middle/upper class Bostonians. The latter two are not mutually exclusive since social varieties are based to a degree on local variety. As Halliday,

MacIntosh and Strevens (1964) suggest:

As one moves along the socio-economic scale, dialectal variety according to region diminishes. Finally at the apex there is no regional differentiation at all, except perhaps for the delicate shades which separate Cambridge and Oxford from each other and from the rest.

Registers are varieties also, but of a different sort than local and social varieties. They are set apart from them by virtue of the immediate circumstances of their use, i.e., the identity of the addressee(s) and other participants in the discourse, the purpose of the communication, etc. The particular register a speaker uses depends on the social situation that triggers its use. Changes in the register used are evidenced in changes in lexicon and in the phonological, syntactic, prosodic and semantic features employed.

There are rules which allow individuals to make judgments of appropriateness of register use and determine when a switch in register is required. These are socially-constrained rules which speaker and addressee have internalized as part of their communicative competence. The exact form and content of these rules is as yet incompletely specified although some things are known (see Ellis and Ure, and section 1.2.2 and 1.3 below).

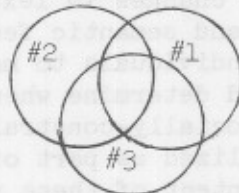
Universality. The notion of register seems to be a universal. Registers are known to operate in all speech communities which have been studied. The specific way in which a given register will operate in a particular speech community is determined by the specific social factors important in the community, how they function and the formal characteristics of the language.

Individual repertoires. Considered together registers cover the total range of language activity. They are not marginal aspects of language or "special" varieties. The entire range of registers reflects the range of social situations in a specific community or culture. They may be used (at least potentially) for specific purposes by all the speakers/writers of a language, thereby cutting across local and social dialectal varieties. Because having control of a register entails having knowledge of and experience with the social situation which governs it, no one individual controls the entire range of registers. No one person has experienced all the social situations in enough depth to learn the entire set. Thus, an individual's repertoire will contain more knowledge about registers in the social domains that he/she is involved directly and less in areas of less involvement. The particular repertoire an individual controls is governed by the common social factors of age, sex, education level, etc.

An individual's competence is more likely to contain rules for a wider range of registers receptively than productively. An adult is likely to control better the production of those registers necessary for functioning in the social domain where he is most active, while

on the receptive side having control over still other registers that he cannot or does not usually have to produce. For example, the male speakers of Berber in northern Africa have full knowledge of registers used productively exclusively by women, although they themselves would never use those registers in any circumstances (James Bynon, personal communication). The case of bilingual speakers is similar since they often have different distributions of register control in the two or three languages known to them. In many instances, this distribution is a function of the situation in which the language(s) was learned and the use(s) to which it is put.

Interrelationship of registers. Registers are interrelated in the sense that there is a great deal of overlap among them. They share many, in fact, most, of their features. One view is to consider the registers in a particular language (regional variety) as largely overlapping sets of features of which some portion of each set is not shared:



This view, however, is similar to the notion of isogloss in discussions of local varieties (dialectal studies) and shares many of the same criticisms. One such criticism is that the presence or absence of features is not absolute. It is more accurate to consider absence/presence as a continuum from greater to lesser frequency of occurrence. In this alternative view, the boundaries between registers are determined quantitatively and set apart from one another by the frequency with which a feature is used. Variables (features) are isolated and a register is defined in relation to other registers quantitatively.

Finally, registers cut across regional varieties to the extent that the social factors defining a particular register are present in the society. For example, it is hypothesized that all speech communities have a register appropriate for use to young children--the baby talk register. This is attested for many different kinds of languages (see 1.3 below) and seems plausible given that all speech communities have children who must acquire language and who must be socialized into the society. This is not to say that the distribution of individuals who control this register will be the same in all communities or that the features defining the register will be identical, but rather that one can isolate some variety in each of these speech communities that is appropriate for use to babies and young children.

1.2.2. "Simplified" speech registers

There is a subgroup of speech registers which have a sufficiently similar set of characteristics that they are sometimes grouped together in discussions on types of speech registers. Ferguson (1971)

has labelled these "simplified" registers. The term is used to refer to conventionalized varieties of speech which are utilized by speakers in situations where the addressee does not have full understanding of the language.

The label "simplified" indicates the process that is claimed to occur when the speaker is placed in a special kind of speech situation. When the level of linguistic competence between speaker and addressee is unequal, the speaker (here assumed to have full linguistic competence) is thought to adjust his linguistic output to a level that is "simpler," "more basic," "more "clarifying." What exactly can be called "simpler" is a matter open to question if one wants to establish strict criteria. However, on a more intuitive level many of the defining features of "simplified" registers do indeed appear to simplify, as will become more evident in the discussion of these characteristics. We know that it is possible for a scientist to take a complex theory, e.g., Einstein's theory of relativity, and "simplify" the content to its basics for the benefit of a less knowledgeable audience (say a freshman college student). In the same way it is possible to "simplify" an utterance in both content and form so that it would be suitable for use to a child with limited knowledge of language and the real world. The same or similar propositional content would be conveyed but in a specific way, a "simplified" form.

It is likely that not all the "simplified" registers or varieties have been isolated or studied at this point, but some of the most commonly discussed examples include: (1) the baby talk register, the variety used primarily in addressing young children, (2) the foreigner talk register (Ferguson, 1971; Ferguson, 1972), (3) the foreign language classroom register (Henzl, 1973), and (4) the register used in speaking to partially deaf persons and other individuals who have physical (or mental) deficiencies that might affect language comprehension and use. At this point it might be useful to review some of the findings of the studies that have investigated these registers for the purpose of comparison and clarification. The information available is limited but sufficient for our purposes. The discussion here will focus on the foreigner talk register and the foreign language classroom register. The baby talk register is discussed in detail in section 1.3 below. The variety used in speaking to partially deaf persons and others has only been mentioned in passing. There has been no attempt to characterize it and therefore it will not be discussed.

The foreigner talk register. Foreigner talk has been studied by Ferguson (1971, 1962) experimentally and through investigation of published sources. The variety has defining phonological, syntactic, and lexical characteristics. Its phonological characteristics include: reduplication, the use of emphatic stress and intonation, slower, louder speech, and more distinct pronunciation. Grammatical characteristics include the omission, expansion and replacement of certain material. Omissions of articles, the copula, conjunctions, subject pronouns, and inflections marking the grammatical categories of case,

person, tense, and number in nouns and verbs are common. Expansions include realization of the subject of imperatives and the addition of tag questions. Replacements affect primarily negative particles, possessive pronouns, and subject pronouns. Lexical characteristics consist of substitutions and analytic paraphrasing, e.g. the use of one day gone and day before this for yesterday, and the use of special vocabulary, e.g. savvy for understand.

Although the characterization of this register is far from complete, this listing indicates the features which exist in speech directed to the addressee who has a poor command of the language of the speaker. It is not difficult to label this speech "simplified." Many of the characteristics serve to increase redundancy (e.g. reduplication) while others eliminate certain surface elements that are unnecessary and redundant for the purpose of expressing basic relationships between objects and events (e.g. omission of person markings on verbs that are redundant with the pronoun).

Foreign language classroom register. The foreign language classroom register also has defining phonological, syntactic and lexical characteristics (Henzl, 1973). The language used by speakers in this study was not English, but Czech. Some of the characteristics overlap with those defining the foreigner talk and baby talk register while others do not. The lack of sufficient data makes it impossible to decide whether these differences are due to the difference between registers or to differences between languages.

Some phonological characteristics of the foreign language classroom register are: few hesitation pauses and slips of the tongue, fully realized vowel quality, fully realized consonant clusters, and more pauses between sentences. Many of these characteristics are probably a direct result of the slower rate of speech that characterizes this register. The grammatical characteristics include the use of shorter sentences, fewer co- and sub-ordinate clauses, and a reduction in the use of inflection in nouns and verbs. In Czech the latter is expressed as an increase in the use of nominative and accusative cases (rather than instrumental) of nouns which in Czech have zero morpheme endings. The lexicon is characterized by the use of a more limited vocabulary, e.g. exclusively using plakat "to cry" rather than all the numerous variants plakala ~ řvala ~ bečela ~ breči, and more specific terms for less specific ones, e.g. "Mr. Smith" and "gentleman" used more than "he," "him".

The purposes of studying the simplified registers have been summarized by Ferguson (1972). He proposes that the study of these registers may clarify the notions of simplification in language and give us a better understanding of the factors that govern language use. In addition, the study of the foreigner talk register may be of value in understanding the process of pidginization (Ferguson, 1971). The study of the baby talk register may advance knowledge of the process of child language development (Ferguson, 1974; Slobin, 1967: 42-5). Thus, the study of these "simplified" registers has value above providing a taxonomic characterization of an aspect of verbal phenomena.

1.3. Baby talk register

1.3.1. Types of studies

The most extensively studied "simplified" register is the baby talk register. The reason for this is probably due to two major factors. For one, the interest in child language development has increased steadily since its beginning in the late fifties. One question which has been raised by many investigators is the effect of the language the child hears on his linguistic development. The second factor to play a role is the strong emphasis placed by some behaviorist theories on the role of language input in language development (Skinner 1957) and the consequent strong negative reaction to that view expressed so eloquently by Chomsky (1959). This ideological controversy brought the issue of the role of language input into the foreground but did not necessarily foster the study of speech directed to children acquiring language, since until the last five years the views of Chomsky prevailed to an overwhelming degree.

Studies of the baby talk register can be grouped into two types. They are sufficiently different in focus and methodology that they deserve to be discussed separately.

Type I studies. Studies of the first type were primarily carried out in the late fifties and throughout the sixties. The study of Arabic baby talk by Ferguson (1956) was one of the earliest studies of this type. The methodology of type I studies is remarkably homogeneous. The investigator gathers information on how adults talk to children in a particular speech community by eliciting from an informant. This information is often supplemented by observations from his own experience and also from information contained in published sources, e.g., novels, diaries, accounts, etc. The source of information is primarily one person, although some investigators consult other informants and occasionally observe a family. Information is usually limited to one or at most two regional varieties. No attempt is usually made to tape record speech.

Type I studies of the baby talk register have been conducted in fifteen languages. These include representatives from a wide range of language families of the world: Arabic (Ferguson, 1956), Berber (Bynon, 1968), Cocopa (Crawford, 1970), Comanche (Casagrande, 1948), English (Ferguson, 1964; Read 1962), Gilyak (Austerlitz, 1956), Greek (Drachman, 1973), Hidatsa (Voegelin and Robinett, 1954), Japanese (Fischer, 1970; Chew, 1969), Kannada (Bhat, 1967), Latvian (Ruke-Dravina, 1961), Romanian (Avran, 1967), Spanish (Ferguson, 1964) and Marathi (Kelkar, 1964). These studies have provided most of the information of a cross-language nature that is available.

Lexical and phonological features are those most commonly discussed in Type I studies, although some passing mention of syntactic devices is often included. The phonological features discussed usually exclude the topics of prosody except for occasional mention of pitch use and tempo (Kelkar, 1964 is an exception here). The focus is on the lexical items that are characteristic in the baby talk register, e.g. in English mommy and daddy for mother and father, and the phonological and morphological processes that distinguish these lexical items.

Type II studies. The studies of the second type came to the fore primarily in the early seventies, when there appeared a number of dissertations on the topic of mothers' speech to children learning language (Broen, 1972; Phillips 1970; Remick, 1971; Snow, 1972). Previous studies of parent-child speech input were conducted at Harvard (Brown, Cazden and Bellugi, 1969; Cazden 1965) and Berkeley (Slobin, 1969; Drach, 1969; Kobashigawa, 1969; Pfuderer 1969), but these were of a more limited nature, i.e. limited in variables studied or number of parent-child pairs observed. Other studies followed, some concentrating in greater detail on specific linguistic features. These are discussed further in 1.3.3. below.

To characterize Type II studies in a more general way, the source information on parent child speech in these studies are tape recordings of mother-child (in some cases non-mothers also) verbal interactions in a variety of structured situations, e.g. playing a game, telling a story, etc. The collection samples are transcribed and certain measures of grammatical complexity and redundancy are obtained. Statistical tests are employed extensively. Most of the measures are of a syntactic nature, although some attention is given to pauses and speech rate (esp. Broen) and most recently to semantic properties (Snow, 1974). All these studies deal with English speakers. No cross-linguistic comparisons are available.

In the two sections that follow (1.3.2 and 1.3.3) I will discuss the major findings of the Type I and Type II studies. Since these studies have been reviewed in detail elsewhere (Type I: Ferguson, 1964, 1974; Type II: Farwell, 1973; Vorster, 1974; Snow, 1974) the discussion will be brief. The discussion is divided into three parts:

- (1) phonological features (other than prosodic) and lexical features of the baby talk register
- (2) syntactic and redundancy features
- (3) prosodic features.

The discussion of prosodic features leads directly into the purpose of the present investigation and the hypotheses to be studied.

1.3.2. Phonological and lexical features

Most of the information on the lexical features and the phonological features (segmental aspects) has been gained in Type I investigations of the baby talk register. Each of these studies has produced some 25 to 60 words which are claimed to make up the lexical items peculiar to the register. The major topics covered by these lexical items are:

- (1) kin names and nicknames, e.g. daddy, mommy,
- (2) body parts and bodily functions, e.g. footsie,
night-night,
- (3) basic qualities ("good", "bad", etc.), e.g. teenie,
itty-bitty "little",
- (4) names of animals and nursery games, e.g. doggie or
bow-wow "dog", piggy-back.

These lexical items have certain foregrounding features. Reduplication is one common feature, as can be seen in items like night-night and bow-wow. Furthermore, diminutive suffixes appear on this set of lexical items in all the languages studied. The most common hypo-choristic suffix in English appears in six of the example items above (doggie, teenie, etc.). In addition these lexical items have a similar canonical form, usually ending with an open syllable. The canonical form varies with the particular language, e.g. CVCV for Berber, (C)VCCV for Japanese, and CVC for Syrian Arabic. The last of these is an example of an exception to the generalization about final open syllables. Phonologically most of these "special" lexical items contain primarily stop consonants, nasals, and a limited selection of vowels. Of the twenty-seven lexical items listed for English by Ferguson (1964) only four contain sounds other than these. The sound "s" appears in footsie and pussy, and the sound "r" appears in burnie and birdie.

Certain phonological processes operate in the baby talk register. They include:

- (1) the deletion or replacement of certain sounds in the adult phonological system, especially r sounds, e.g. English rabbit becomes wabbit (replacement), English drink becomes dink (deletion).
- (2) the loss of certain phonological distinctions present in language directed to adults, e.g. Berber--the neutralization of vowel length distinction.
- (3) distinct nasal assimilation, e.g. Gilyak damk "hard" becomes ama.
- (4) replacement of velars by apicals, e.g. English tum on for come on.
- (5) simplification of consonant cluster, e.g. English tummy for stomach.
- (6) certain interchanges between sibilants, affricates and stops, e.g. Marathi--affricates replaced by stops.
- (7) loss of unstressed syllables, e.g. Spanish tines for calcentines.

Many of these phonological processes move in the direction of less markedness and toward conforming with the phonological rules of children's speech at early stages of language development. This applies to the feature of reduplication as well. An example of the "less marked" rule is the replacement of r in rabbit with the semivowel w (wabbit). A counterexample to this rule, however, is the replacement of s by ch in Spanish, e.g. Spanish becho for beso.

1.3.3. Syntactic complexity and redundancy features.

A number of measures of syntactic complexity and redundancy show that speech to the young child is syntactically simpler and more redundant than speech to adults. This is manifested in a variety of ways. The following are the major characteristics of speech directed

to a child listener as opposed to an adult listener.

Co-ordination and subordination. Speech addressed to the child contains few or no embedded or conjoined clauses (Drach, 1969; Sachs, Brown, and Salerno, 1972; Snow, 1972; Phillips, 1973). The proportionate number of relative clauses, complements, and subordinate clauses is much lower in adult speech when the speaker is addressing a child. For example, Drach found that subordinate clauses are ten times more frequent in speech addressed to the adult than to the child.

Inflections. Speech addressed to a child listener contains fewer grammatical inflections than speech addressed to the adult. Snow (1972) found fewer inflections for both nouns and verbs. Many Type I studies also mention this as characteristic of other languages too.

Repetitions. There is a high incidence of repetitions in speech addressed to the child. One study (Kobashigawa, 1969) found that 35 percent of all utterances directed to the child listener are repetitions of some type (exact, partial, paraphrases). Not all sentence types are subject to repetitions equally. About 60 percent of imperatives, 25 percent of questions, and 15 percent of statements are repeated. Snow (1972) found that most repetitions are partial repetitions usually produced immediately after the full sentence form. Noun phrases and prepositional phrases were the parts of the sentence that were most frequently repeated. In addition, a number of utterances directed to the child listener are paraphrases. These may be viewed as a type of repetition.

Type/token ratio. The type/token ratio of speech directed to the child listener is smaller than that in speech directed to the adult (Broen, 1972; Drach, 1969; Phillips, 1970 and 1973; Remick, 1971). A small type/token ratio indicates a restricted and repetitious vocabulary. This finding and the immediately previous one point to the great amount of redundancy in speech addressed to children. This applies especially to children under the age of six years.

Number and length of sentences. Speakers use a greater number of sentences when addressing a child listener and these sentences are much shorter (Drach, 1969; Remick, 1971; Sachs, Brown, and Salerno, 1972; Snow, 1973). These two findings are undoubtedly related. Sentences used to the adult are an average of 2 1/2 times as long as those used to the child listener and about 10 percent of all utterances directed to the child consist of single word utterances (Broen, 1972). It is not surprising that more sentences are used to the child than to the adult in the same time period.

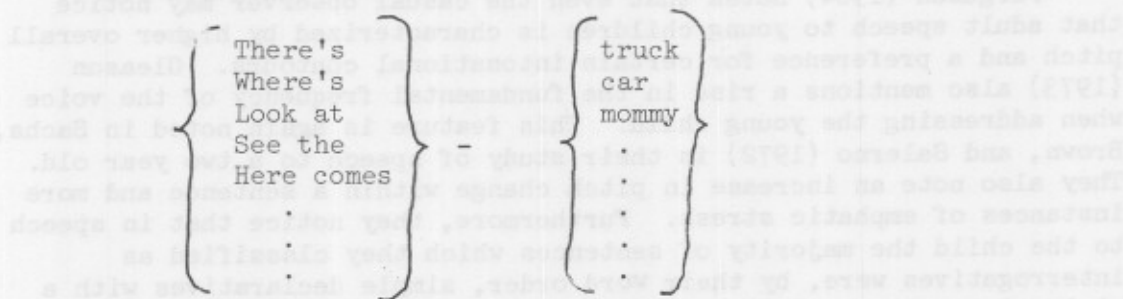
Verbal routines. One common verbal routine that introduces vocabulary items in a set frame has been noted by several investigators (Broen, 1972; Ferguson, Peizer, and Weeks, 1973). An example dialogue illustrates its use:

MOTHER TO CHILD: Look, Tommy.
Look at the truck.
See the truck.
Where's the truck.

CHILD: Truck.

MOTHER TO CHILD: Yes, truck.
Here comes the truck.

The underlined sentences indicate use of the set frame. In this example the frame sentences were produced with the same rhythm and heavy primary stress on the last word (truck). The frames may be schematized as follows:



Use of such frames has been found to be as common as five instances of the same frame produced within a five minute period (Broen, 1972).

Question and imperative sentence forms. The proportion of question and imperative sentence forms as opposed to declarative forms, increase in speech directed to the child listener. This feature is one of the best documented findings available (Blount, 1972; Broen, 1972; Drach, 1969; Pfuderer, 1969; Remick, 1971; Sachs, Brown, and Salerno, 1972; Snow, 1972). Questions, especially, are more frequent with estimates varying from 35 percent to 65 percent as compared to 70 percent declarative forms in speech to the adult listener. Imperatives account for about 30 percent of the utterances directed to the child. Ervin-Tripp (1970) suggests that a high degree of interrogative forms are the result of the greater need of adults to ask for feedback from the child.

Personal pronouns. The distribution of personal pronouns in speech to the child listener has been extensively investigated by Willis (1974). One of the most frequent findings in this area is the decrease in use of the third person pronouns in speech directed to the child (Broen, 1972; Snow, 1972; and a number of Type I studies).

The list presented here is not exhaustive but rather covers those findings that form the broad outlines of the syntactic complexity and redundancy features of mothers' speech to children.

1.3.4. Prosodic features.

Up to this point the discussion has focused on the particular lexical and syntactic characteristics of adult speech addressed to

young children. However, some investigators of the baby talk register have also noted the characteristic use of prosodic features, i.e. "features whose arrangement in contrastive patterns in the time dimension is not restricted to single segments" (Lehiste, 1970: 3). These features are referred to as pitch, stress, and quantity and may function on a paralinguistic level (Markel, 1965). Some acoustic correlates of these are duration, fundamental frequency (F_0) and intensity. These features play a role in all languages but in various ways. There are observations on the use of these features in the baby talk register for several languages.¹ The importance of such information for the student of language development is discussed in Crystal (1973). The focus of the discussion that follows is on the information presently available on the use of such features in the baby talk register by American English speakers.

Ferguson (1964) notes that even the casual observer may notice that adult speech to young children is characterized by higher overall pitch and a preference for certain intonational contours. Gleason (1973) also mentions a rise in the fundamental frequency of the voice when addressing the young child. This feature is again noted in Sachs, Brown, and Salerno (1972) in their study of speech to a two year old. They also note an increase in pitch change within a sentence and more instances of emphatic stress. Furthermore, they notice that in speech to the child the majority of sentences which they classified as interrogatives were, by their word order, simple declaratives with a rising intonation. Only a small portion of the "interrogative" sentences contained question words or had inverted word order. Sachs et al. suggest that rising sentential intonation may signal something other than "question" in speech to the child, and that the rising intonation may be a special kind of pitch change.

These same characteristics are mentioned by Grewel (1959) who astutely added a number of other features to the list on the basis of his own casual observations of adult speech to children. He notes a higher overall pitch and, in simple sentences, a rising intonation at the end of the sentence. He observes, further, that longer sentences are divided up in sections with each section having its own completed rising or falling contour. Frequent successive repetitions of the same contour are also noted. Speech to the young child is slower in tempo with obvious (prolonged--OKG) pauses between words, word groups, and particularly between sentences.

Grewel makes some further comments which though suggestive, are ambiguously stated and difficult to interpret. He asserts that "the dynamic accent in speaking to babies is strikingly diminished as compared to speaking to adults" and also that "when a dynamic accent is used, it is as if it were compensated by a prolongation of the stressed word." (Grewel, 1959: 196). As an example of the latter he offers the sentence: "No, that we ca__'nt do!" The term dynamic accent is not defined and could be interpreted as meaning any of several different things. From the example above, it seems at least likely that Grewel is using the term "dynamic accent" to mean emphatic stress. In any case, the prolonged duration of stressed words also

seems to be a probable characteristic of speech directed to the child.

Remick (1971) reports the only empirical evidence on prosodic feature characteristics. She studied the speech of ten mothers to their children (ages 1;4 to 2;5). She calculated both median fundamental frequency and frequency range from narrow band spectrograms for a subsample of sentences from each subject. The spectrographic analysis was run on fourteen to seventeen utterances per subject from each of two distinct speech situations: (1) speech directed to an adult, and (2) speech directed to the child addressee. Her finding was that only the mothers of the youngest children used a higher median fundamental frequency and a greater frequency range when addressing the child. The speech of mothers whose children had begun to acquire language showed "a dramatic restriction in both median and range" (Remick, 1971: 32).

Several methodological inadequacies in this study lead us to question the data on fundamental frequency and range as well as the conclusions drawn from them. Some of these inadequacies are acknowledged by the investigator. First, since the recordings were made in the subjects' home, the quality of the recordings was in all likelihood poor. The choice of sentences for spectrographic analysis was thus biased toward those with a more favorable signal to noise ratio. The investigator reports that only a limited number of readings could be obtained even from the measurable sentences. Second, in reporting the findings Remick gives no account of the procedure used for making measurements. This leaves open the question of how certain decisions were made, decisions that could have profoundly affected the values obtained. Third, there was no attempt to match sentences measured from each of the two situations in terms of their composition. A number of investigations (Peterson and Barney, 1952; House and Fairbanks, 1953; Lehiste and Peterson, 1961) have shown that vowels have intrinsic pitch, i.e. there is a connection between vowel quality and the relative height of the average fundamental frequency associated with it: higher vowels have higher fundamental frequency. Therefore, in any attempt to compare the fundamental frequency in two or more situations or even across subjects within a single situation, it is necessary to obtain measurements on the same verbal material. If the phonetic composition of the samples varies greatly, the differences observed may only reflect a difference in the composition of the two samples. Finally, no statistical tests were run on the frequency data. Therefore, it is not clear that the observed differences between speech to the child and speech to the adult were significant.

One final observation about prosodic features in speech to young children is that the use of such characteristics seems to diminish or disappear in most contexts by the time that child addressee is four to five years of age. This has been noted in the case of higher fundamental frequency (Gleason, 1973; Grewel, 1959). The situation for the other characteristics is as yet unknown.

1.4. Hypotheses

Research on the use of prosodic features in the baby talk register is a broad area of study. Although passing comments have been made on

this topic (section 1.3.4 above), practically no systematic investigation has been carried out. The present study represents an initial inquiry into this area, and will therefore, be limited to the study of only a few selected aspects of the problem.

The hypotheses to be tested in this study are:

Hypothesis 1: The use of prosodic features in adult speech directed to young children differs systematically from the use of such features in speech to other adults. (That such differences are evidenced in the corresponding acoustic parameters is presupposed.)

Hypothesis 1a: The fundamental frequency of speech to the child is higher than that of speech to the adult.

Hypothesis 1b: The frequency range of speech to the child is greater than that of speech to the adult.

Hypothesis 1c: The duration of syllable nuclei of words receiving primary stress is greater in speech to the child than the adult.

Hypothesis 1d: The use of a rising pitch terminal in declarative sentences is more frequent in speech to the child than in speech to the adult.

Hypothesis 1e: In speech to the child the basic sentence unit is more frequently segmented into shorter sections than in speech to the adult.

Hypothesis 2: The degree and character of all the above differences (hypotheses 1a-1e) vary as a function of the relative age of the child addressee. Generally the older the child is, the closer the use of prosodic features will approach the pattern in speech to the adult.

All of the above hypotheses are held to apply under experimental conditions where the speech context remains constant and only the addressees are varied.

Footnote to Chapter 1

1. An extensive commentary on the particular use of prosodic features in the baby talk register in Marathi is found in Kelkar (1964). Passing references to one or two features are also found in reports on this register in other languages (see bibliography).

CHAPTER 2: METHOD

2.1. Subjects

The subjects were twenty-four women college graduates under thirty-five years of age residing in the predominantly white, upper middle class suburban community surrounding Stanford University. The subjects were native speakers of American English with minimal or no knowledge of a foreign language. Their speech was devoid of any discernible speech disfluency. All the women had lived in California for at least one year.

The women were contacted initially by telephone and asked to participate in the study. They were told that the study was concerned with the ability of children of different ages to pay attention in a set of common, everyday situations. Each subject agreed to participate in two sessions. The subject was told that in the first session she would be familiarized with a set of situations and that in a second session she would interact with her own children in these situations. None of the subjects seemed aware that their own behavior rather than the behavior of the children was of primary interest.

Twelve of the women had a child in the 1;10 - 2;6 age range (mean age - 2;3). The other twelve women had a child in the 5;1 - 5;7 range (mean age - 5;4). There were an equal number of male and female children in each age group. Table 1 lists the age and sex of the children that participated in the study.

Table 1
Age and sex of children participating with their mothers in the study.

Subject	Age of Child	Sex of Child	Subject	Age of Child	Sex of Child
C101	2;6	M	C201	5;3	M
C102	2;1	M	C202	5;1	M
C103	1;10	M	C203	5;6	M
C104	2;5	M	C204	5;2	M
C105	1;10	M	C205	5;6	M
C106	2;2	M	C206	5;2	M
C107	2;6	F	C207	5;5	F
C108	2;5	F	C208	5;3	F
C109	2;6	F	C209	5;2	F
C110	2;0	F	C210	5;4	F
C111	2;3	F	C211	5;7	F
C112	2;6	F	C212	5;7	F
\bar{X} age = 2;3 Age range = 1;10 - 2;6			\bar{X} age = 5;4 Age range = 5;1 - 5;7		

2.2. Procedure

2.2.1. Testing sessions

Each adult took part in two testing sessions, and performed three verbal tasks in each session. In the first session the subject

directed her speech to another adult (adult listener session). The adult listener in all cases was the investigator. In the second the subject directed her speech to her own child (child listener session). In a counterbalanced design the adult-child session should precede the adult-adult session for half the subjects. However, it was highly desirable to obtain speech samples in both sessions that were as natural as possible. The possibility of justifying an adult-adult session following the adult-child session without revealing the purpose of the experiment seemed remote. On the other hand, all subjects accepted the request for the adult-adult session before the adult-child session for the purpose of familiarizing them with the materials.

2.2.2. Verbal tasks

The subjects performed three tasks in each testing session: (1) a picture task, in which the subject told a story about the persons and events depicted in each of a series of pictures, (2) a story reading task, in which the subject read a short descriptive passage, (3) a puzzle task, in which the subject gave a series of instructions on how to solve a puzzle.

The order in which the tasks were performed by each adult was different in each testing session. The adults in each group (C1 and C2) were randomly assigned to one of the ordering sequences in the adult-listener session and to a different one in the child listener session. For example, if a subject in the adult-listener session performed the tasks in the order: reading task, puzzle task, picture task, then the order in the child-listener session was different, e.g. picture task, story reading task, puzzle task.

2.2.2.1. Picture Task

Five colored pictures depicting situations thought to be of interest to children were chosen from several magazines. The pictures showed a boy eating a hamburger, some boys dressed up like Indians sitting around a campfire, a family on a picnic, a family doing household chores, and a little girl and her mother baking. Each picture was mounted on a 9" x 11" piece of colored cardboard. A short declarative sentence related to the events depicted in the picture was written below each photograph. Detailed descriptions of each picture and the accompanying sentences are given in Table 2. The subjects were asked to make up a short story to go along with each picture. They were told to incorporate the exact wording of the sentence accompanying the picture into their story. The instructions for the adult-listener session and the child-listener session went as follows:

8.2.1. Testing sessions
 8.2.1.1. Procedure
 Each adult took part in two testing sessions, and performed three verbal tasks in each session. In the first session the subject

Table 2
Description of materials in Picture Task.

PICTURE #1

Description: A small boy is sitting on a porch step. In his lap is a plate with a large hamburger. A large shaggy dog is standing next to the boy. The dog is eating out of his food dish.

Sentence: They are both hungry.

PICTURE #2

Description: Three small boys each wearing an Indian headdress and wrapped up in a blanket are sitting outdoors around a campfire. In the kettle boiling over the fire is some soup.

Sentence: It is cold.

PICTURE #3

Description: A family is having a picnic outdoors. The mother, father, and little boy are sitting on a blanket spread out on a grassy area. There is food in front of them. It is a sunny day.

Sentence: They are glad that it didn't rain.

PICTURE #4

Description: A large family (mother, father, numerous children, grandfather) is standing in front of their house. Each member of the family is holding some object indicating that they are preparing to do some household chore (the girl is carrying a large basket of laundry, the father has a bucket in one hand, etc.).

Sentence: Everybody is doing his chores.

PICTURE #5

Description: A woman and a young girl are in the kitchen. They are both wearing aprons. The woman is helping the little girl place some dough into a baking pan.

Sentence: Next time the girl will do it herself.

Adult-adult session

"I am going to show you pictures of some familiar events. Look at the pictures one at a time and tell a short story about the people and events you see. Each of your stories should be about four or five sentences long. One of these sentences should be the one below the picture. Please give this sentence exactly as it appears there."

Adult-child session

"Here are the pictures you saw last time. (Hand pictures to subject). Today I would like you to tell _____ (child's name) a short story about the people and events pictures in

them. Your stories should be about four or five sentences long. One of these sentences should be the one listed below the picture. Please give that sentence exactly as it appears."

After the instructions were given, the subject was given the opportunity to ask questions about the task. Then the subject began the task. Since the task was both simple and straightforward, few problems arose. Occasionally in the adult-listener session, a subject would either forget to use the sentence accompanying the picture or would rephrase the sentence. In these cases the investigator commented on this and stressed the importance of including that sentence in the story exactly as it appeared. The subject was then asked to repeat her story, making the correction.

2.2.2.2. Puzzle task

In the puzzle task the subjects were asked to give verbal instructions to the listener on how to solve a puzzle. The puzzle was a small wooden object in the shape of a barrel. Solving the puzzle involved taking the barrel to pieces. The barrel could be taken apart only by pushing and removing certain pieces in a specified order. The subjects were told that the barrel could be taken apart and were given the necessary instructions. They were to give these instructions one by one to the addressee so that he/she could disassemble the puzzle. The pieces composing the barrel were color coded to facilitate this process. The instructions were as follows:

1. Push in the green square.
2. Take out the piece.
3. Push in the red piece.
4. Take out the piece.
5. Push in the blue piece.
6. Take out the orange piece.
7. Take out the purple piece.
8. Take out the brown pieces.

The five year old children, by and large, had no problem identifying the correct pieces and following the instructions. Some of the two-year olds, however, either did not recognize some of the color terms used in the instructions or confused them. In these cases the mother was instructed that she could assist the child only after each instruction was presented as it appeared in (1)-(8) above. The instructions to the subjects on this task were as follows:

Adult-adult session

"This object is a barrel. (Show object to subject). It is a puzzle and can be disassembled into a number of pieces. The colored pieces must be removed in a certain order. This is a list of the steps a person would have to follow in order

to solve the puzzle. (Give subject list of instructions.) Give me the instructions one by one as they appear on the list and I will take the barrel apart. That way you can see yourself how the puzzle works."

Adult-child session

"Here is the barrel puzzle you saw me work last time. (Give subject puzzle.) This time _____ (child's name) will be taking apart the puzzle. You are to give _____ (child's name) the instructions on this list exactly as they are written there, in the order in which they appear. (Give list of instructions to subject.) You can assist _____ (child's name) if he/she has difficulty following any instruction but please give the instructions first as it appears on the list."

On a few occasions the pieces of the puzzle were moved incorrectly by a child so that the disassembly of the puzzle could not be continued. In these cases the investigator who was observing the proceedings interjected some instructions to the mother so that the situation could be righted and the process could continue.

2.2.2.3. Story Reading Task

A short passage about rainbows accompanied by a picture depicting a rainbow and other items mentioned in the passage (sun, raindrops, etc.) was presented to the subject. The passage was the first paragraph of the Rainbow Passage (Fairbanks, 1940):

When the sunlight strikes raindrops in the air, they act like a prism and form a rainbow. The rainbow is a division of white light into many beautiful colors. These take the shape of a long round arch with its path high above, and its two ends apparently beyond the horizon. There is, according to legend, a boiling pot of gold at one end. People look, but no one ever finds it. When a man looks for something beyond his reach, his friends say he is looking for the pot of gold at the end of the rainbow.

The subject's task was to read the passage out loud. In the adult-listener session the subject simply read the passage out loud in the presence of the investigator. In the adult-child session the subject read the passage to the child and could interact verbally with the child if she wished. This was included because pilot subjects had expressed the desire to have such an option. They felt that such interaction more closely approximated the manner in which they read stories to their children at home.

The instructions to the subjects were as follows:

Adult-adult session

"Accompanying this picture is a short story about rainbows. (Give picture to subject.) So that you will have read the passage through once, please read the passage out loud now."

Adult-child session

"This is the story about rainbows which you read last time. I would like you to read the story to _____ (child's name) as you might read a story to him/her at home. If you like, you may point out objects in the picture which are mentioned in the story or ask _____ (child's name) questions about them. If, while you are reading, you are interrupted in the middle of a sentence, please start again at the beginning of that sentence and proceed from there."

The Rainbow Passage was chosen for several reasons. First, the subject matter was thought to be of interest to children even though some of the sentences were quite complex. And, second, Horii (1972) has shown that there is a high correlation (+.98) between average fundamental frequency measurements for the second sentence and average fundamental frequency measurements for the rest of the passage. This suggests that measurements made in the second sentences would be very close to the values obtained if every sentence in the passage was measured. Thus, it would be possible to measure a small sample (one utterance per subject) and these measurements would be generalizable to the entire passage.

2.2.3. Interviews

An interview was conducted with each subject upon completion of the verbal tasks in the child-listener session. The interview was conducted in the presence of the child, who was occupied with some toys provided by the investigator. The purpose of the interview was to determine whether the subjects in the study were aware that they modified their speech when addressing their children, especially whether they noticed any prosodic modifications. If they were aware of such changes, what kinds of modifications did they notice?

The questions included in the interview appear below. The rather broad (imprecise) terms "same sort of voice" was used in an attempt to elicit comments from the subjects on the prosodic aspects of their speech. The probes were also structured to steer the subjects' responses in this direction. A direct question containing the term "prosodic" was ruled out because of the difficulty in explaining the meaning of this highly specialized terminology.

Interview Questions:

- (1) Did you notice whether you spoke with the same sort of voice when addressing _____ (child's name) as you did when you addressed me in the previous session?

If "yes" answer: Was your voice different? What sort of differences did you notice? PROBE (if necessary):

- Was it higher or lower in pitch? Softer or louder? Was there more or less fluctuation in the pitch of your voice?
- (2) Have you noticed any such differences on other occasions or in other situations? On what occasions? In what circumstances?
 - (3) Have you ever observed that other parents speak in a different sort of voice to their children than they do to an adult? Who was it? What differences did you notice? On what occasions did this occur? What were the circumstances involved?
 - (4) If you want to get your child's attention, or get him/her involved in conversation, how would you do this? Specifically, would you change your voice in a particular way? What way(s)?

After the interview was completed, the subjects were informed of the actual purpose of the study and of the general hypotheses and predictions advanced by the investigator. They were encouraged to pose questions regarding any aspect of the study. All the subjects concurred that they had not been aware that their own speech rather than the child's behavior was the primary focus of the study.

2.2.4. Data Collection

2.2.4.1. Taping of sessions

All the testing sessions were recorded on tape in their entirety. The subjects were told that all the sessions were being recorded, including the adult sessions. The reason given for taping all the sessions was that the study had many participants and that the investigator would therefore be unable to rely on her memory alone in reviewing the various sessions. Most of the subjects accepted this reasoning without any further questioning. A few questioned the necessity of taping the adult-listener session when the child was not present. They were told that the investigator would review the adult-listener session prior to the child-listener session to determine if the tasks were performed in the manner expected. This was, in fact, done in a number of cases. No subject persisted in questioning the reasons for taping the testing sessions once these reasons were given.

2.2.4.2. Physical setting

All the testing sessions took place in an acoustically treated room, ordinarily used for recording by the university foreign language laboratory. Its dimensions were 6' by 9'. This room was selected in order to ensure tape recordings of a high quality. The room was on the inside of the building and had no windows. The only outlet was to a short hallway that led to the inside corridor of the building. The experimental room was equipped with special features to minimize the effect of sound waves reflecting off the flat surfaces--floor-to-ceiling buffers lined a portion of each wall, the floor was covered with a thick wall-to-wall carpet, and acoustical tile covered the

ceiling. The 2" thick solid wood door was insulated on all sides for a tight seal with the door frame. The room was illuminated by an overhead incandescent light to avoid the low frequency noise ("hum") often emitted by fluorescent bulbs.

The tape recorder was placed on a small table in one corner of the room. Next to the table was a chair where the investigator sat during the child-listening session. Two large, brightly-colored floor pillows were placed in the middle of the room parallel to one another and approximately 18" apart. During the testing sessions, the speaker (the subject) and the addressee (the investigator in the adult-listener session; the child in the child-listener session) each sat on one of the pillows facing one another.

The exact placement of the speaker and the microphone was determined by a preliminary test. The investigator read aloud a short excerpt from the Rainbow Passage in a normal voice several times. These readings were recorded in the experimental room on the same equipment used during the testing sessions. The speaker's position in the room was varied, as well as the left or right deviation from the straight ahead position of the speaker's head in relation to the microphone. Subsequently these recordings were judged by the investigator and one other person for quality of recording.¹ A minimum of distortion and fluctuation as a result of head movement was sought. On the basis of these sample recordings, the two judges determined the location of the speaker within the physical layout of the room that yielded the best recording. In the study the subject always sat in this location.

2.2.4.3. Instrumentation

The tape recordings were made on a Revox A77 tape recorder using a Sony Electret condenser microphone and Scotch 176 tape (1/4" x 1200'). The tape recorder was calibrated just prior to the beginning of the study to give a flat frequency response of ± 2 db over the range of 50-10,000 Hz. The calibration was checked once approximately half-way through the study, and a second time upon completion of the study. The tape recorder performed reliably. The microphone was attached to a lavalliere placed around the subject's neck. The microphone hung approximately 10" from the subject's mouth.

2.2.5. Data Processing

2.2.5.1. Selection of utterances

A total speech sample of 30-40 minutes was obtained for each subject (both sessions combined). Since it is extremely costly in terms of time and data processing equipment to analyze such an enormous amount of data, a subsample of the utterances was chosen for perceptual and acoustical analysis. The utterances selected for analysis were the sentences provided in the picture task (5 sentences), the instruction sentences from the puzzle task (8 sentences), and the second sentence from the reading task (1 sentence). This yielded a sample of 14 sentences from each of 48 testing sessions. A total of 672 sentences were analyzed.

These particular sentences were selected in order to make inter-session and inter-subject comparisons on samples in which the lexical items were the same. With lexical content held constant, the analysis could focus on the properties that were of interest in the study.

2.5.2.2. Acoustic Analysis

Instrumentation. The utterances selected for analysis were dubbed from the original recordings using a duplicate Revox A77 tape recorder. The dubbing was done by means of a machine-to-machine patch cord leading from the output jack of machine No. 2. These dubbed utterances were then processed on the Pitch Extractor in the Phonetics Laboratory at the University of California at Berkeley (Krones, Ms.).² This Pitch Extractor produces a display indicating the fundamental frequency (F_0) of the voiced portions of utterances.

The Pitch Extractor is an analog device operating in real time so that the pitch contours may be recorded. Its output is a voltage that ranges from -10 to 0 volts which varies according to the frequency of the input signal. The Oscillomink is used to record this voltage and display the pitch contour on a roll of calibrated paper which is approximately 5 1/2" wide. The top half of the display is used to record the pitch contours while the bottom half is used to record amplitude (voltage). The Pitch Extractor was used in conjunction with a Transpitchmeter which supplied the input filter and calibration tones.

The Pitch Extractor can be adjusted for the frequency range that it will measure. The maximum frequency to be measured can be varied from 100, 150, 200, 300, 400, or 600 Hz. The lower limit to be measured is then represented by a percentage of the maximum frequency-- 25 percent, 50 percent, 75 percent. In the processing of utterances it is first necessary to select a range of measurement for each set of sentences. Once this range is determined, frequencies above the maximum are clipped off automatically and frequencies below the lower limit indicating lack of phonation, are not calibrated for frequency. A range must be set for each set of utterances that will encompass the range of frequencies represented in the voiced portions of the utterances. Calibration tones are used to indicate the frequency represented at particular points in the display within the established range.

Initially, a small subset of the utterances selected for analysis were processed on the Pitch Extractor. Twenty-four utterances from three speakers (2 C1 subjects and 1 C2 subject) were chosen for this preliminary analysis. There were an equal number of utterances from the adult listener session and the child listener session for each speaker. These utterances were selected as representative of the entire sample in terms of quality of recording and pitch range.

The purpose of this preliminary analysis was to determine whether the dubbed tape recordings produced satisfactory displays and whether it would be necessary to set different calibrations for the same subject when analyzing utterances from the adult-listener and child-listener

sessions. The resultant displays were judged to be satisfactory for measurement purposes. However, there were sufficient differences between the two sessions, as well as among subjects, to suggest the need for frequent recalibrations from session to session.

The total set of utterances was then processed on the Pitch Extractor. The machine was optimally calibrated for the utterances produced by each subject in each testing session. Calibration tones were used to record the frequencies represented on the displays for each new calibration. The displays were produced at what was judged to be an optimum rate--100 mm/sec. The displays produced for each subject were inspected for instances where the frequency was outside the maximum or minimum of the optimal frequency range. Utterances in which this had occurred were processed again with a new calibration.

2.2.5.3. Perceptual Analysis

The utterances selected for analysis were transcribed by the investigator. The total transcription consisted of a broad phonetic transcription of the segmental portion of each utterance in IPA notation, and a transcription of the accentual pattern. Four levels of stress were marked: Stress 1 (primary stress), Stress 2 (secondary stress), Stress 3 (tertiary stress), and Stress 4 (unstressed).

2.2.5.4. Measurement of fundamental frequency and duration

Measurements of fundamental frequency and duration were made for each syllable nucleus in each utterance. The following information was recorded: (a) the fundamental frequency at the beginning, peak, and end point of the syllable nucleus, (b) the location of the peak within the syllable nucleus, (c) the duration of the syllable nucleus and (d) the intensity at the peak of the syllable nucleus.

Clear, plastic templates were constructed from the calibration sheets for measuring fundamental frequency. A separate template was made for each calibration. Horizontal lines representing calibration tones were drawn parallel to the base line appearing in both the calibrations and the frequency display recordings. The lines represented specific frequencies on the Hertz (Hz) scale, e.g. 210 Hz, 220 Hz, etc. The interval between each pair of lines was 10 Hz. The template was superimposed on the fundamental frequency display for each syllable nucleus and the frequency of the beginning, peak, and end point was determined. If the point occurred exactly in the middle of an interval, it was assigned a value half-way between the two values defining the interval. For example, a point in the middle of the interval between 210 Hz and 220 Hz was assigned the value 215 Hz. Points which occurred within the interval but which were closer to one of the two values were assigned that value. Measurement error was estimated at ± 5 Hz.

In most cases the syllable nucleus corresponded to a separate syllable in the utterance. However, in certain cases the boundaries between two syllables were not well enough defined on the frequency display and the two syllables were considered, for measurement purposes, as a single syllable nucleus. For example, this was the case for the

words in the in some speakers' renditions of the sentence "Push in the [ɪnə] green square."

Occasionally the end point of the last syllable nucleus in an utterance was impossible to measure because the subject's voice exceeded the limitations of instrumentation at the lower values. It was expected that such instances of unmeasurable phonation would arise. The frequency with which this occurred varied with the subject and type of session. In no case did this occur, in more than 5 percent of the sentences in one session. When it did occur, the lowest observed frequency value for that subject was assigned.

Duration was measured from the beginning point to the end point of each syllable nucleus. A transparent metric ruler was used. Since the display was produced at the rate of 100 mm/sec, one mm on the ruler was equal to .01 seconds. The location of the frequency peak in the syllable nucleus was also recorded.

The intensity curves for each utterance were recorded directly below the frequency display. Intensity was not calibrated to an absolute scale. Therefore, the values for intensity corresponding to the peak of a syllable nucleus were recorded in millimeters. This made it possible to compare the intensity at the peak of one syllable nucleus with the intensity at the peak of another syllable nucleus.

Intra-observer reliability was obtained for measurements of fundamental frequency and duration. A 10 percent sample of sentences was chosen randomly for remeasurement. Remeasurement of duration yielded values identical to the initial measurement. Therefore, no statistical analysis to determine reliability was performed. Remeasurement of fundamental frequency did yield slightly different results. The correlation between the initial fundamental frequency measurement and the remeasurement values was +.97. The intra-observer reliability was considered acceptable.

Footnote to Chapter 2.

1. I am grateful to Dr. Dorothy Huntington for giving her expert advice on this matter.
2. I thank Jean Marie Hombert for his assistance in processing this material.

CHAPTER 3: RESULTS

3.1. Fundamental frequency

The average fundamental frequency was computed for each subject for each session from the speech samples that were measured. These data are shown in Tables 3 and 4.

Table 3
Comparison of fundamental frequency and frequency range data for C1 and C2 subjects.

	C1 subjects (N=12)		C2 subjects (N=12)	
	Adult listener	Child listener	Adult listener	Child listener
Mean fundamental frequency (Hz)	197.6	267.3	202.8	206.4
Mean fundamental frequency (Hz)	43.2	48.4	43.6	43.9
Total range (st)	10.5	19.2	10.9	12.6

Hz = hertz
st = semitones above the zero frequency level of 16.35 Hz.

Table 4
Average fundamental frequency for C1 and C2 subjects by session.

Subject	C1 subjects				C2 subjects				
	Adult listener (Hz)	Child listener (st)	Adult listener (Hz)	Child listener (st)	Adult listener (Hz)	Child listener (st)	Adult listener (Hz)	Child listener (st)	
C101	201	43.4	238	46.4	C201	185	42.0	193	42.7
C102	181	41.6	240	46.5	C202	202	43.5	208	44.0
C103	188	42.3	274	48.8	C203	215	44.6	216	44.7
C104	178	41.3	211	44.3	C204	188	42.3	197	43.1
C105	195	42.9	269	48.5	C205	199	43.3	198	43.2
C106	215	44.6	300	50.4	C206	239	46.4	241	46.6
C107	189	42.4	276	48.8	C207	198	43.2	197	43.1
C108	241	46.6	302	50.5	C208	186	42.1	188	42.3
C109	178	41.3	272	48.7	C209	206	43.9	205	43.8
C110	193	42.7	296	50.1	C210	206	43.9	223	45.3
C111	207	43.9	274	48.8	C211	234	46.1	237	46.3
C112	205	43.8	256	47.6	C212	176	41.2	174	40.9

Hz = hertz

st = semitones above the zero frequency level of 16.35 Hz

The mean fundamental frequency for the C1 group was 197.6 Hz in the adult-listener session and for the C2 group it was 202.8 Hz. These figures are within the expected range of values for female speakers (Linke, 1953; Peterson and Barney, 1952; Snidecor, 1951). The variations from speaker to speaker result primarily from differences in the size of the vocal folds. Other factors that can affect the values obtained are the nature of the verbal material spoken by the subject and the particular measure that is used to compute the values (mean versus median). The difference between the means for C1 and C2 subjects in speaking to an adult was not significant ($t = .75$; $df = 11$; $t_{(.05)} = 2.20$).

The mean fundamental frequency for the C1 and C2 subjects in the child-listener sessions were 267.3 Hz and 206.4 Hz respectively. For the C2 subjects, where the listener was a 5 year old, the frequency level in speech to the child was not very different from the level in speech to the adult. The difference between these two means is not significant ($t=2.07$; $df=11$; $t_{(.05)}=2.20$). The small differences between the two types of sessions is evident in the graphic representation in Figure 1. In only one C2 subject was the difference between speech to the adult and speech to the child more than one semitone (1.4 sts).

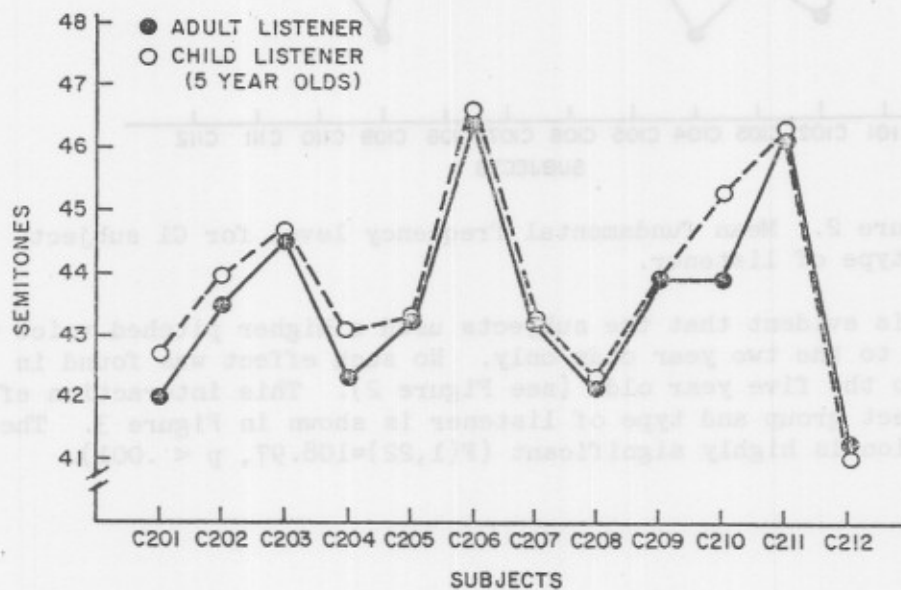


Figure 1. Mean fundamental frequency level for C2 subjects by type of listener.

The difference between the means for the child listener and adult listener conditions for the C1 subjects is quite large--197.6 Hz versus 267.3 Hz. This difference is highly significant ($t=11.55$; $df=11$; $t_{(.001)}=3.11$). For all C1 subjects, the fundamental frequency

level in speech to the child was considerably higher than in speech to the adult. The differences between the two types of sessions ranged from 3.0 semitones up to 7.4 semitones. These large differences are quite evident by visual inspection of Figure 2.

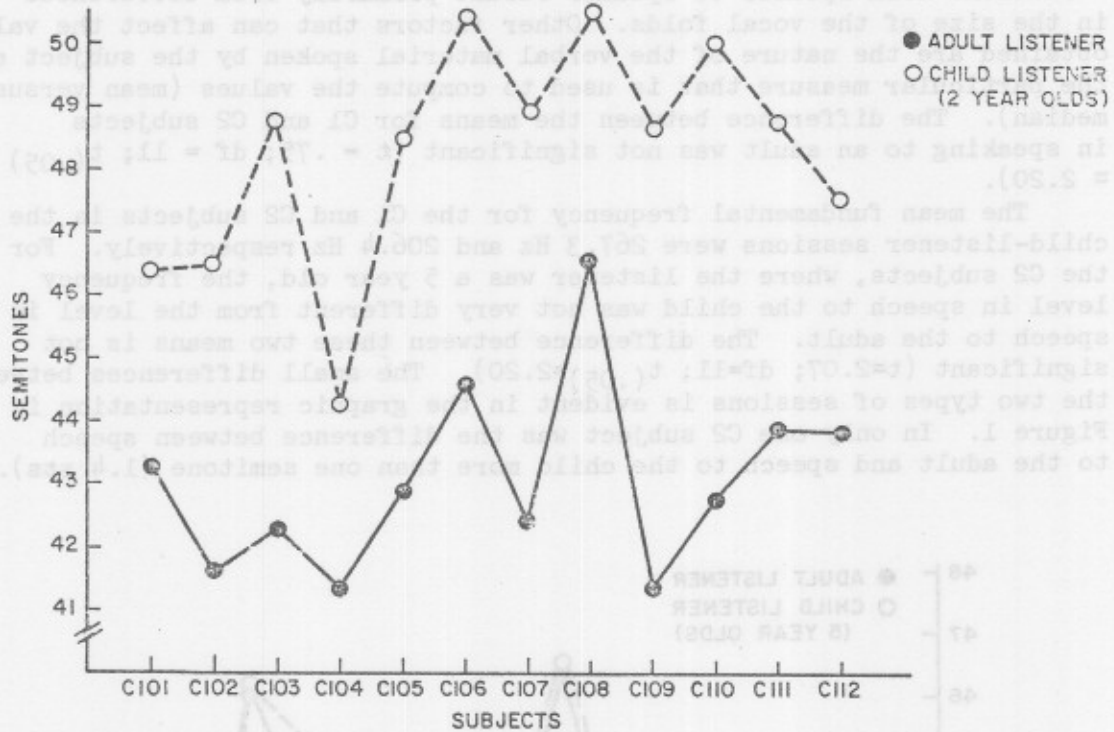


Figure 2. Mean fundamental frequency level for C1 subjects by type of listener.

It is evident that the subjects used a higher pitched voice when speaking to the two year olds only. No such effect was found in speech to the five year olds (see Figure 2). This interaction effect for subject group and type of listener is shown in Figure 3. The interaction is highly significant ($F(1,22)=108.97, p < .001$).

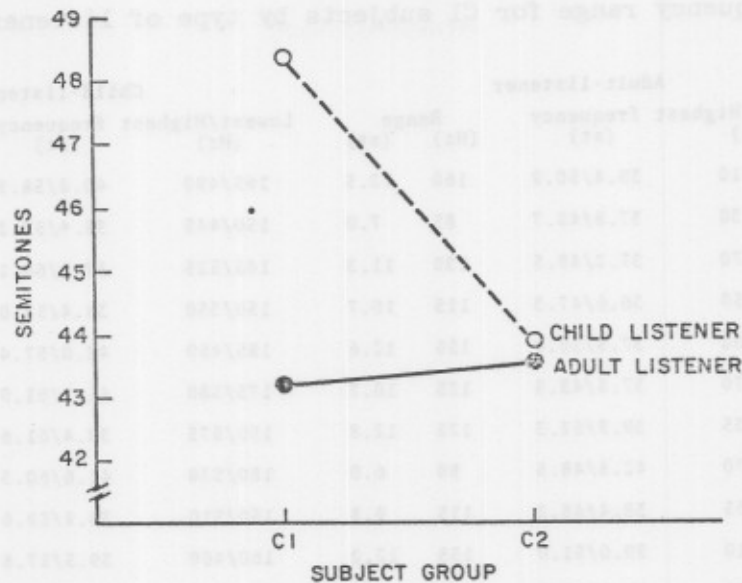


Figure 3. Mean fundamental frequency level by subject group and type of listener.

3.2. Frequency range.

The frequency range for each subject was determined for each session separately. The range was defined by the lowest and highest frequency produced in the session, shown by the figures in Tables 5 and 6. The ranges represent the lowest and highest frequencies actually employed by the subjects in their speech during the session rather than their phonational frequency range. The latter term refers to vocal frequencies ranging from the lowest sustainable tone in the modal register to the highest sustainable tone in falsetto. The ranges for the C1 and C2 subjects in the adult-listener sessions extend from 75 and 80 Hz at the low end to 160 and 185 Hz at the high end. These figures represent a span of approximately 1/2-1 octave (Figures 4 and 5). This range span corresponds well to findings for female speakers by other investigators (Duffy, 1958; Linke, 1953). The mean ranges for the C1 and C2 subjects in the adult listener sessions are quite similar--10.5 and 10.9 semitones. This difference was not significant ($t=.61$; $df=11$; $t(.05)=2.20$).

The frequency ranges of the C1 and C2 subjects in the child-listener sessions were greater than those in the adult-listener sessions. The smallest and largest spans were 200 and 425 Hz among the C1 subjects and 125 and 250 Hz among the C2 subjects. For a number of C1 subjects, the ranges approach a two octave span. The differences between the frequency ranges used to adult-listener versus child-listener for the C1 subjects was highly significant ($t=9.48$; $df=11$; $t(.001)=3.11$). This difference in the range in speech addressed to an adult-listener versus a child-listener was also significant for the C2 subjects ($t=3.376$; $df=22$; $t(.001)=3.11$).

Table 5
Frequency range for C1 subjects by type of listener.

	Adult-listener				Child-listener			
	Lowest/Highest (Hz)	frequency (st)	Range (Hz) (st)		Lowest/Highest (Hz)	frequency (st)	Range (Hz) (st)	
C101	150/310	38.4/50.9	160	12.5	165/490	40.0/58.9	325	18.9
C102	145/230	37.8/45.7	85	7.9	150/445	38.4/57.2	295	18.8
C103	140/270	37.2/48.5	130	11.3	145/525	37.8/60.1	380	22.5
C104	135/250	36.6/47.3	115	10.7	150/350	38.4/53.0	200	14.6
C105	145/300	37.8/50.4	155	12.6	185/450	42.0/57.4	265	15.4
C106	145/270	37.8/48.5	125	10.7	175/580	41.1/61.9	405	20.8
C107	160/335	39.5/52.3	175	12.8	150/575	38.4/61.6	425	23.2
C108	190/270	42.5/48.5	80	6.0	180/530	41.6/60.5	350	18.7
C109	150/265	38.4/48.2	115	9.8	150/510	38.4/59.6	360	21.2
C110	155/310	39.0/51.0	155	12.0	160/460	39.5/57.8	300	18.3
C111	160/260	39.5/47.9	100	8.4	170/510	40.6/59.6	340	19.0
C112	130/245	36.0/46.9	115	10.9	150/450	38.4/57.4	300	19.0
			\bar{X}	10.5			\bar{X}	19.2

Table 6
Frequency range for C2 subjects by type of listener.

	Adult-listener				Child-listener			
	Lowest/Highest (Hz)	frequency (st)	Range (Hz) (st)		Lowest/Highest (Hz)	frequency (st)	Range (Hz) (st)	
C201	130/240	35.9/46.5	110	10.6	150/310	38.4/50.9	160	12.5
C202	140/280	37.2/49.2	140	12.0	140/310	37.2/50.9	170	13.7
C203	145/310	37.8/50.9	165	13.1	150/300	38.4/50.4	150	12.0
C204	145/265	37.8/48.2	120	10.4	140/265	37.2/48.2	125	11.0
C205	160/240	39.5/46.5	80	7.0	140/265	37.2/48.2	125	11.0
C206	155/340	38.9/52.5	185	13.6	150/400	38.4/55.3	250	16.9
C207	150/315	38.4/51.2	165	12.8	140/310	37.2/50.9	179	13.7
C208	145/240	37.8/46.5	95	8.7	150/275	38.4/48.9	125	10.5
C209	155/320	38.9/51.5	155	12.6	150/310	38.4/50.9	160	12.5
C210	150/280	38.4/49.2	130	19.8	150/315	38.4/51.2	165	12.8
C211	180/290	41.5/49.8	110	8.3	150/330	38.4/52.0	180	13.6
C212	140/260	37.2/47.9	120	10.7	130/250	35.9/47.2	120	11.3
			\bar{X}	10.9			\bar{X}	12.6

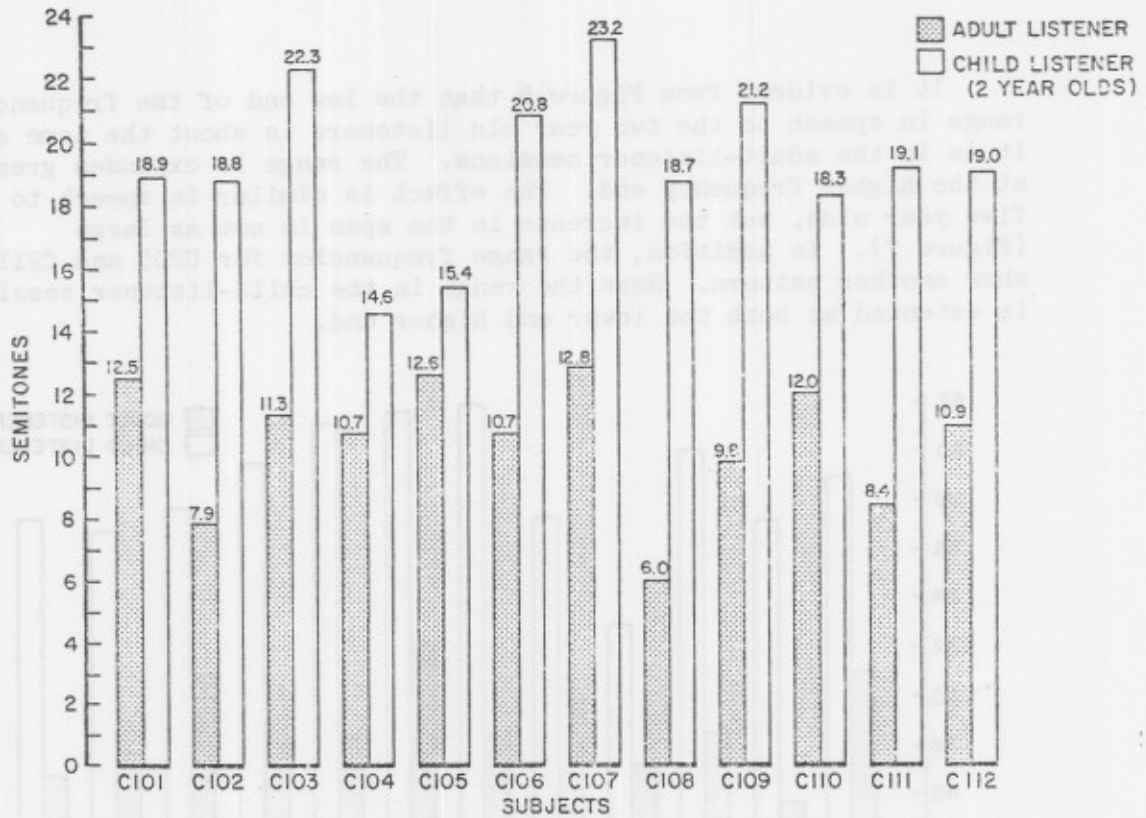


Figure 4. Comparison of frequency ranges (in semitones) for C1 subjects by type of listener.

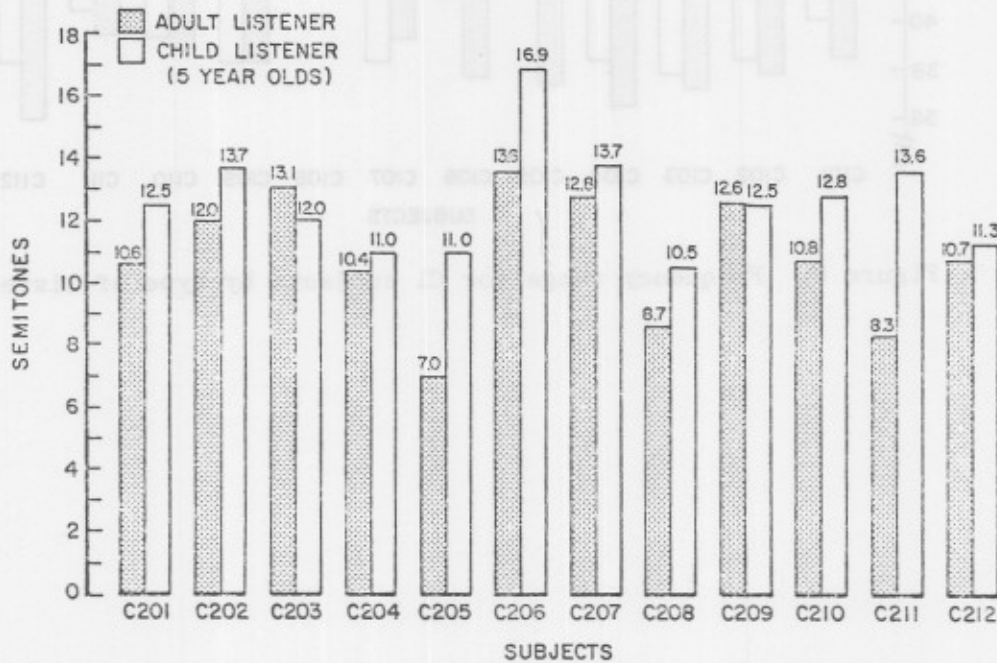


Figure 5. Comparison of frequency ranges (in semitones) for C2 subjects by type of listener.

It is evident from Figure 6 that the low end of the frequency range in speech to the two year old listeners is about the same as it is in the adult-listener sessions. The range is expanded greatly at the higher frequency end. The effect is similar in speech to the five year olds, but the increase in the span is not as large (Figure 7). In addition, the range frequencies for C205 and C211 show another pattern. Here the range in the child-listener session is extended at both the lower and higher end.

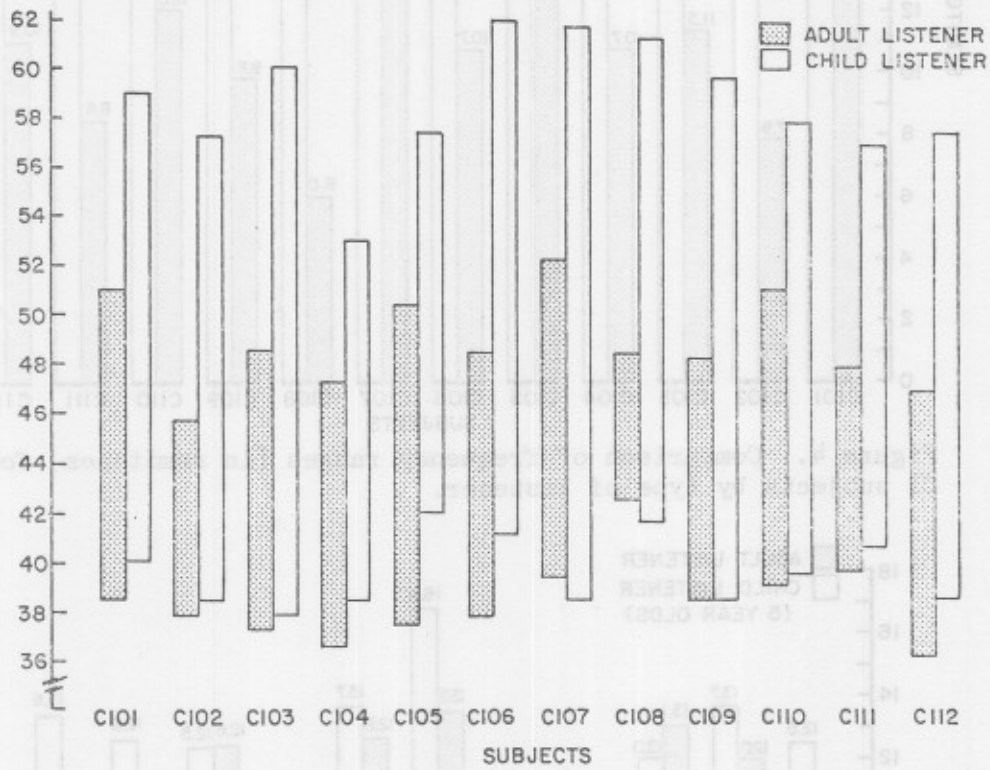


Figure 6. Frequency range for C1 subjects by type of listener.

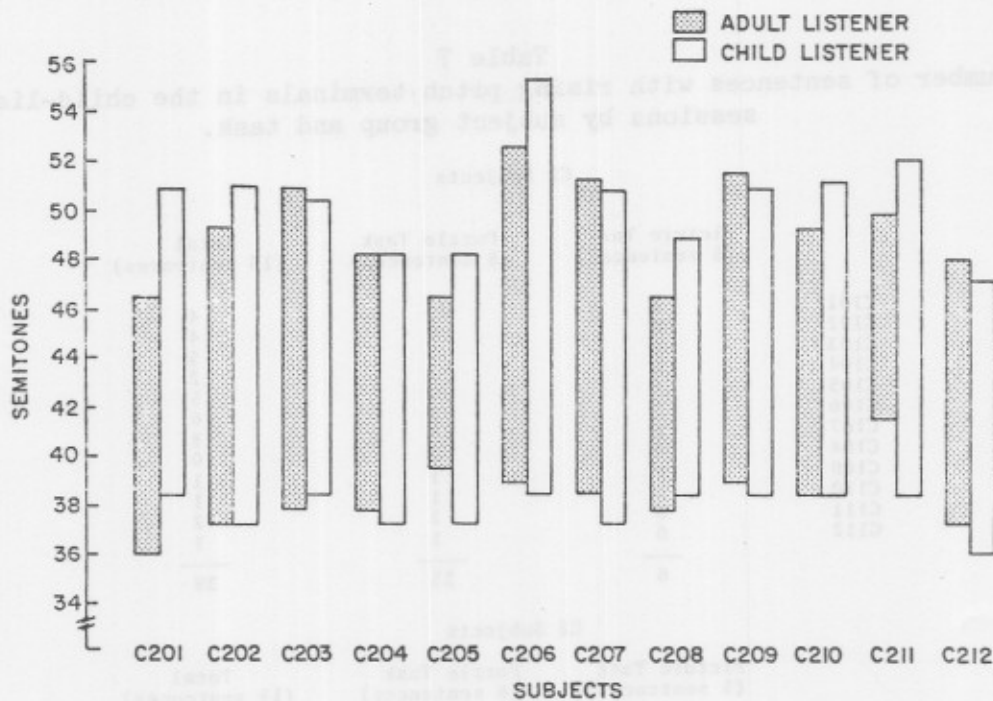


Figure 7. Frequency range for C2 subjects by type of listener.

3.3. Sentence final pitch terminals.

The sentences sampled from the Picture task were declarative and imperative and therefore we would expect these sentences to have a falling final pitch terminal. All the sentences were produced with a falling terminal by the subjects in the adult-listener sessions. However, in the child-listener sessions this was not the case. A prominent feature in speech to the two year olds (C1 subjects, child-listener session) was a rising final pitch terminal in the Puzzle task sentences and occasionally in the Picture task sentences. The number of sentences with a rising final pitch contour for each task is shown in Table 7. Twenty five percent of the sentences spoken by the C1 subjects ended with a rising terminal. All but one subject (C108) used the rising terminal in at least one sentence during these two tasks. This result is highly significant (Fischer sign test, $B=11$, $b.0032$, 12 , $1/2$). Furthermore, most of the sentences with rising pitch terminals (85 percent) were produced in the Puzzle task and were therefore all imperative sentences.

Table 7
Number of sentences with rising pitch terminals in the child-listener sessions by subject group and task.

C1 Subjects

	Picture Task (5 sentences)	Puzzle Task (8 sentences)	Total (13 sentences)
C101	1	3	4
C102	0	4	4
C103	0	5	5
C104	0	2	2
C105	3	2	5
C106	1	5	6
C107	1	7	8
C108	0	0	0
C109	0	1	1
C110	0	1	1
C111	0	2	2
C112	0	1	1
	<hr/> 6	<hr/> 33	<hr/> 39

C2 Subjects

	Picture Task (5 sentences)	Puzzle Task (8 sentences)	Total (13 sentences)
C201	0	0	0
C202	0	0	0
C203	0	0	0
C204	0	1	1
C205	1	1	2
C206	0	1	1
C207	0	1	1
C208	0	0	0
C209	1	1	2
C210	0	1	1
C211	0	0	0
C212	0	1	1
	<hr/> 2	<hr/> 7	<hr/> 9

A rising pitch terminal was also present in speech to the five year olds but was much less common. Nine percent of the sentences spoken by the C2 subjects in the child-listener sessions ended with a rising terminal. Seven out of the twelve C2 subjects used the rising terminal in at least one sentence in the two tasks but no subject used it more than twice. This is not statistically significant (Fischer sign test, $B=7$, $b_{.3872}$, 12, 1/2). Here again most of the sentences with rising terminals were produced in the Puzzle task. Thus, we find that when the two year old child is the addressee some of the sentences are produced with a rising terminal, even though they are statements and, surprisingly, imperatives. Ordinarily, in adult speech the rising terminal is restricted to questions. The rising terminal was also used occasionally in speech to the five year olds.

3.4. Use of whispering.

An unexpected finding was the preponderance of whispering used by the C1 subjects in the child-listener session. Whispering never

occurred for either C1 or C2 subjects in the adult-listener sessions. Only two C2 subjects (C206 and C208) used whispering in the child-listener session, and then each used it in only one of the sampled sentences. However, nine out of the twelve C1 subjects used whispering in at least one sentence when speaking to the two year olds. The number of sentences (out of 13) in which whispering was used by each subject was: 1 sentence (C106, C112), 2 sentences (C105, C110, C111), 3 sentences (C101, C102, C109) and 4 sentences (C107). In only one case was an entire sentence whispered. Most often the last half of the sentence contained the whispered syllables. A check of the complete transcripts of the child-listener session for the C1 subjects revealed that the use of whispering was not restricted to the sub-sample chosen for acoustic and perceptual analysis but was evident throughout the entire session; in some cases, the use of whispering was more extensive in the unanalyzed portions of speech. In the case of the three C1 subjects that did not use whispering in the sentences analyzed, all of them made at least some use of whispering at other points in the child-listener session. Of the C2 subjects, only the two subjects mentioned above made any use of whispering in the child-listener session. The use of whispering by and large seems to be restricted to speech to two year olds.

3.5. Duration

Two content words from each of six sentences in the Puzzle task were chosen for the comparison of average duration of syllables between the adult-listener and child-listener sessions. The six sentences (1, 3, 5, 6, 7, 8--see 2.2.2) contain both a verb (push in/take out) and a color term (green/red/blue/orange/purple/brown). The average duration of the syllable nucleus (in msec) was computed for each subject by session. Verb and color terms were tabulated separately for Tables 8 and 9. In computing the average durations for the verbs, the two word sequence (e.g. push in) was considered as one and the durations were added together. In the case of purple, the only color term pronounced with two syllables, the two syllables were also considered as one item. The color term orange was always pronounced as one syllable, e.g. [arndʒ].

The results for the verbs and color terms will be discussed separately. For the verbs, the difference between the means for the C1 and C2 subjects in the adult-listener session was highly significant (510.91 and 523.08 respectively). This difference was not significant

Table 8
Average duration (msec) of verbs in Puzzle Task

	C1 Ss		C2 Ss		
	Adult listener	Child listener	Adult listener	Child listener	
C101	263.33	301.67	C201	260.00	236.00
C102	190.00	243.33	C202	286.67	275.00
C103	193.33	183.33	C203	205.00	228.00
C104	225.00	266.67	C204	266.67	266.33
C105	265.00	296.67	C205	190.00	201.67
C106	205.00	238.33	C206	195.00	186.67
C107	176.67	203.33	C207	218.33	211.67
C108	191.67	318.33	C208	198.33	198.33
C109	148.33	181.67	C209	166.67	167.67
C110	210.00	208.33	C210	223.33	236.67
C111	193.33	230.00	C211	198.33	261.67
C112	270.00	315.00	C212	268.33	275.00
	$\bar{X} = 210.97$	$\bar{X} = 248.89$		$\bar{X} = 223.06$	$\bar{X} = 228.72$

Table 9
Average duration (msec) of color terms in Puzzle Task

	C1 Ss		C2 Ss		
	Adult listener	Child listener	Adult listener	Child listener	
C101	238.33	310.00	C201	246.67	251.67
C102	210.00	243.33	C202	210.00	243.33
C103	183.33	223.33	C203	226.67	321.67
C104	190.00	192.00	C204	255.67	296.67
C105	285.00	355.00	C205	181.67	205.00
C106	200.00	218.33	C206	208.33	260.00
C107	205.00	271.67	C207	201.67	230.00
C108	205.00	301.67	C208	205.00	240.00
C109	188.33	218.33	C209	141.67	187.50
C110	205.00	238.33	C210	251.67	271.67
C111	205.00	361.67	C211	200.00	296.67
C112	255.00	398.33	C212	273.33	355.00
	$\bar{X} = 214.16$	$\bar{X} = 277.67$		$\bar{X} = 216.86$	$\bar{X} = 263.26$

The results for the verbs and color terms will be discussed separately. For the verbs, the difference between the means for the C1 and C2 subjects in the adult-listener session was fairly small (210.97 and 223.06 respectively). This difference was not significant

(Wilcoxon rank sum test, $W^* = 0.87$, $z(.05) = 1.645$). However, for ten out of twelve C1 subjects the average duration of the verbs was higher in the child-listener session than in the adult-listener session. Only six of the C2 subjects showed a similar difference. The difference between the durations in the adult-listener and child-listener session for the C1 subjects was significant (Wilcoxon signed rank sum test, $I^* = 2.30$, $z(.05) = 1.645$) and this is reflected in the difference between the means for the two sessions--210.97 and 248.89. The difference between the means for the two sessions for C2 subjects was smaller (22.306 and 228.72) and was not significant [Wilcoxon signed rank test, $B = 6$, (.6128, 12, 1/20)].

The results for the color terms are slightly different. The difference between the means for the C1 and C2 groups in the adult-listener session (214.16 and 216.86) was very small and not statistically significant (Wilcoxon rank sum test, $W^* = 0.52$, $z(.05) = 1.645$). However, for all the subjects there was an increase in average duration of color terms in the child-listener session. The difference between the means for the two sessions for both C1 and C2 subjects reflects this fact (C1 subjects--214.16 versus 227.67; C2 subjects--217.86 versus 263.26). Both are significant (Wilcoxon summed rank test, $I^* = 3.45$, $z(.001) = 3.09$).

One factor which has not yet been considered in this analysis is the perception of differences in duration. It is not the case that a one millisecond difference between two stimuli will be perceived and this should be considered in interpreting the results. In fact, the just noticeable difference (jnd) for duration increases as the duration of the standard stimuli increases. There are several studies on the perception of durations but they do not seem to agree on the size of the jnd for different duration values (Stott, 1935; Henry, 1948; Ruhm et al. 1966). However, if we approximate conservatively from the available information (i.e. that 30 msec is the jnd for durations of 150-200 msec, 35-40 msec for durations 200-250 msec, and 45 msec for durations for 250-300 msec), we should have enough information to correct for the potential effect of the perceptual factor in evaluating differences in duration between the adult-listener and child-listener sessions. For C1 subjects we find that the difference between average duration values for the adult-listener and child-listener sessions are larger than the perceptual threshold. This is the case in all instances for the verbs. For the color terms, this is the case for all subjects except C104 and C106. Therefore, taking the perception factor into consideration does not change the results for the C1 subjects. The original results also hold for the C2 subjects. For verbs, the difference between the adult-listener and child-listener sessions is not significant. Taking the perceptual factor into consideration reduces the number of changes from adult-listener to child-listener to one subject (C212). For color terms, the perceptual factor likewise does not alter the results since only two subjects (C201 and C209) are affected. Thus, we find that the duration of verbs and color terms is significantly longer in sentences spoken to the two year olds. Only

the duration of the color terms is longer in sentences spoken to the five year old children.

3.6. Distribution of primary stress

One result of the perceptual analysis was the finding that primary stress placement was different in speech directed to the child-listener than to the adult.¹ This occurred only in speech directed to the two year olds. The difference observed was the appearance of two primary stressed syllables in a sentence which ordinarily, in adult-adult communication, would contain only one primary stress. The sentences in which this phenomenon was noticed were the six sentences of the form "Push in _____" and "Take out _____" in the Puzzle Task.

As shown in Table 10, there were only scattered instances of use of double primary stress in speech directed to the adult listener. Five out of 144 sentence samples fall into this category, amounting to approximately three percent of the sentences directed to the adult listener in the Puzzle task. Only three of the C1 subjects (C101, C105, C112) and two of the C2 subjects (C202, C207) showed use of double primary stress. This stress distribution appeared in only one sentence for each of these subjects.

Table 10
Number of instances of two primary stresses per sentence in six sentences from the Puzzle task by type of listener.

	C1 Subjects		C2 Subjects	
	Adult listener	Child listener	Adult listener	Child listener
C101	1	2	C201	-
C102	-	1	C202	1
C103	-	-	C203	-
C104	-	3	C204	-
C105	1	3	C205	-
C106	-	2	C206	-
C107	-	1	C207	1
C108	-	1	C208	-
C109	-	-	C209	-
C110	-	1	C210	-
C111	-	1	C211	-
C112	-	3	C212	-
	3/12 Ss	10/11 Ss	2/12 Ss	1/12 Ss

The situation is somewhat the same in speech directed to the five year olds. Only one C2 subject (C202) used double primary stress in the child-listener session. On the other hand, ten of the C1 subjects used double primary stress in at least one sentence (Fischer Sign test, $B = 10$, $b_{.0193}$, 12, 1/2). Of these, three subjects used double primary stress in each of three sentences from the Puzzle task, two subjects in

two sentences each, and five subjects in one sentence each. This represents a total of 18 out of 72 sentences or about 25 percent of the Puzzle task sentences directed to the two year old child listener. Only two C1 subjects (C103, C109) did not exhibit this characteristic in their speech.

It seems evident, therefore, that the assignment of more than one primary stress to the short and simple sentences of the Puzzle task occurs almost exclusively in the case where the sentences are directed to the two year old. Only scattered instances occur in speech directed to the adult listener and the five year old.

3.7. Summary of results.

Six major analyses were performed on samples of speech directed to adult listeners and to child listeners. Speech to the two year olds differed on the six analyses from speech to the adult listeners. Only some of these differences were found between speech directed to the five year old child listener and the adult listener.

The results indicate that:

(a) The average fundamental pitch of the speaker's voice is higher to the two year old than to the adult. This was not the case for speech to the five year old.

(b) The frequency range of the speaker's voice is greater to the two year old and to the five year old in comparison with the speech range to the adult listener. The expansion occurs at the high end of the range.

(c) Speech to the two year old contains many instances of rising sentence final pitch terminals in sentences where the grammatical form would normally dictate a falling pitch, e.t. imperatives. This feature is absent from speech directed to the adult listener and to the five year olds.

(d) Whispered or partially whispered words appear in speech directed to the two year olds. This characteristic is absent from speech directed to the adult listener and to the five year old.

(e) The duration of certain content words in selected sentences is prolonged in speech to the child listener as compared to that of the adult listener. In the Puzzle task sentences, the verbs and the color terms had longer durations in speech to the two year olds than to the adults. Only the durations of the color terms were so affected in speech to the five year old.

(f) Speech directed to the two year old contains many cases of more than one instance of primary stress assigned within a sentence unit. This feature is absent in speech directed to the five year old and to the adult listener.

Footnote to Chapter 3.

1. Two transcribers independently transcribed primary stress placement in these sentences. A correlation of +.92 was found between the transcribers.

CHAPTER 4: DISCUSSION

4.1. Characteristics of the baby-talk register: functions

It has been shown in this and other studies that speech directed to young children differs in systematic ways from speech directed to an adult. Given that these findings are reliable and that these differences do exist, the question that arises is what functions these particular features serve. Of what potential use is this "special" speech to the child learning language? I would suggest that the various features of the baby talk register can serve at least two functions--an analytic function and a social function.

A child learning language is constantly analyzing the speech that he hears as a means for learning the rules of his language. He then uses these rules to effectively communicate with those around him. This analytic endeavor is no doubt an enormous task. Although some theorists, notably Chomsky (1965), have suggested that language input of almost any kind is sufficient for language learning, this seems unlikely. Speech directed to children is rich in redundancy, repetition and other devices which multiply the linguistic information available in the utterance. A plausible hypothesis is that characteristics found to be prevalent in the baby talk register have a distinct function in the child's analytic endeavor. They assist the child in delimiting sentences, words and other syntactic constituents and possibly do more. Among the characteristics of this register that have such an analytic function are some of the aspects of prosody discussed in the previous chapter. This is because the prosodic features of speech are the primary means by which a speaker organizes units above the level of the phonological segment into groups.

Knowing specific linguistic rules, however, such as those governing proper word order is not enough. To communicate effectively the child must know, among other things, the rules for how to engage in a verbal exchange with another person. And, in order to interact with a child, it is necessary to gain the child's attention for conversation. Getting the initial attention of your interlocutor in a conversation is a primary prerequisite to beginning a communicative exchange. Keeping the attention of your interlocutor is necessary for the maintenance of communication. It is hypothesized that some features of the baby-talk register serve a social function, i.e. to initiate and maintain communication between adult and child (and likewise between an older child and a young child).

Prosodic features can cue the child to pay attention and listen to the speech of the person attempting to communicate with him. Other verbal means used include frequent repetitions of the child's name. The features which serve a social function complement those features which assist the child's analytic endeavor, since a child must attend to a particular set of speech in order to utilize whatever analytic cues are provided.

In the following two sections the findings of the present study will be considered in terms of the above two functions. It will

become apparent that some features may play a dual role, that is, simultaneously serve a social and an analytic function. Other features seem to serve predominately one function or the other.

4.1.1. Prosodic features: social function

The higher pitched voice used by the subjects in this study to the two-year-olds can be viewed as serving primarily a social function. The higher pitch is quite unique to this function. It may in fact be the most salient characteristic that serves to mark and thus set apart the baby talk register from other registers. Even the most casual of observers seems to notice it. An utterance spoken with a higher pitched voice marks that message as intended for the child listener. The message may in other respects be 'tailored' to the language abilities of the child. A message so marked prosodically is foregrounded against the background of adult-adult communication.

The question arises of whether the higher pitch level is in some way more salient to the child. There seems to be little if any evidence that the young child's hearing apparatus is more sensitive to higher frequencies than lower ones. It is the case, however, that in general higher frequencies are more audible than lower frequencies, given the same intensity level. The higher pitch level is, of course, closer to the pitch of the child's own voice. One can only speculate as to whether this factor plays a role. In any event, it is at least plausible that a higher pitch level serves a social function by regulating communication with the child. It attracts the child's attention to verbal material directed to him.

The expanded pitch range observed in the speech of the adult subjects to both groups of child listeners (two year olds and five year olds) also has a similar function. The extension of the range was primarily in the upward direction. The presence of high pitch peaks in utterances intended for the child listener, may be salient cues that mark particular sections of a speaker's speech and therefore make them stand out. The finding of higher pitch in speech to the child listener should be considered with the understanding that such speech is not characterized by high levels of pitch in every syllable nucleus. It is rather the case that the peaks in a sentence unit that do appear are in many cases exaggerated in comparison with speech to adults.

The use of whispering also may be considered as an example of the social function of prosodic characteristics. Whispering, in fact, is very closely allied with the extension of the range capabilities of speech. Whereas the range is expanded at the higher end by the presence of higher syllable peaks, an extension at the lower end of the range may result in the voice going into whisper. There is ample evidence from languages using tone that when a speaker produces an exaggerated rendition of a low tone, a whisper may result. In Serbo-Croatian, for example, Ivić and Lehiste (1969) observed that the voices of speakers who exaggerated the low-to-high tone at the end of an utterance went into whisper on the low tone portion of the

utterance. In some African languages when there is a lowering of tone at the end of questions, whisper often appears (Will Leben, personal communication). Thus it seems that the expansion of range in the baby talk register occurs at both the high and low end of the voice range. Both the high pitched syllable and the whispered syllable stand out and perhaps have attention getting properties.

Finally, the preponderance of rising sentence final terminals in speech to the child listener may serve a social function--to regulate conversation between adult and child. The predominance of rising terminals may cue the child as to when he is expected to respond, since the question is the grammatical form most often associated with a rising terminal, and questions normally demand an answer. Also, it has been noted that sustained or rising pitch in place of terminal falls is generally used to indicate "unfinished business." (Bolinger, 1961).

It is not uncommon to observe an adult asking a child listener a question and then answering the question if the child does not respond to complete the exchange. The completeness of the question/answer sequence in terms of a communication unit is best seen if one thinks of the question forming the first half of a contour (ending with a rising pitch terminal) and the answer continuing the contour and ending with a falling pitch terminal signals completion of the contour and simultaneously the completion of the exchange. The presence of many rising pitch terminals may serve, then, not only to regulate the conversation between adult and child but also to keep the child's attention. One must pay attention in a conversation in order to know when it is one's turn to speak.

Thus, the higher pitch level, the expanded pitch range, the use of whispering, and the predominance of rising pitch terminals in adult speech directed to a child listener can be interpreted as serving a social function.

4.1.2. Prosodic features: Analytic function

Some of the prosodic characteristics discussed in the previous chapter seem to serve an analytic function. One feature which plays a dual role is the preponderance of rising pitch terminals. These may be used to cue the child to the location of sentence boundaries. The fact that a high pitch is attained at the end of the sentence (the boundary) is significant because the high pitch would tend to accentuate the termination of the sentence by the speaker. Furthermore, the rising pitch terminals were associated with sentences which by grammatical form were imperatives and not interrogatives. It is unlikely that the adults were using the imperatives as questions since the context of the sentences was the administration of instructions to a task. Making the imperatives into questions would indicate that the speaker was unsure of the instructions. This never occurred even in the adult-adult sessions when the subject was completely unfamiliar with the task and the instructions she was to administer. It is more likely that the rising pitch terminals on sentences of the imperative

form functioned as a signal both to regulate the verbal exchange (social function) and to mark the sentence boundary (analytic function).

The longer durations of certain words (one or both content words in the sentences studied) can also be seen as potentially having an analytic function. Duration is an important correlate of stress, although there is no direct, one-to-one relationship between the duration of a syllable and the degree of stress it carries. For the following discussion, the situation will be somewhat simplified by disregarding other factors involved.

In speech to the five year olds the durations of the color terms were significantly greater than those directed to the adult listener. The extension of duration in color terms can be viewed as a way to supplement the function of contrastive stress on the unit. By prolonging the duration of the syllable nucleus of red in Push in the red piece the speaker implies with greater force the propositions "not the yellow piece, not the blue piece."

In speech to the two year olds, the duration of both the color terms and the verbs were greater than to the adult listener, indicating both emphatic stress on the verbs and contrastive stress on the color terms. This may be the "key" words in the sentence. These were the only words the child needed to understand in order to carry out the command correctly. For example, in the case of the first sentence Push in the green square, the listener had only to attend and understand the words push and green to correctly complete the demanded action. The word square is redundant here since there were no other green pieces.

The longer durations of verbs and color terms to the two year olds no doubt contributed greatly to the perception of two primary stresses in many sentences directed to them. When two primary stresses were transcribed, they were marked as falling on the verb and color term of the sentence. Aside from the above mentioned function (to indicate keywords), two primary stresses may serve to divide up a sentence perceptually into smaller units. The adult thereby segments the sentence into pieces he/she thinks are of adequate size for the child to process easily. The same sentence which, when directed to the adult, would normally contain only one primary stress, would be divided into multiple units for the child. Furthermore, it is interesting that the chunks that sentences such as Push in the green square are divided into are the major constituents of the sentence. By this division, the adult may be providing the child with important information about constituent structure. A look at more sentences with different structures would be necessary before anything more than a tentative statement on this point could be made. However, my initial inspection of all the sentences contained in the Rainbow Passage, as read to the child listener, confirms this position. In reading to the child listener, (to the two year olds in particular) the longer sentences are divided up prosodically (here with respect only to stress) into smaller units. These smaller units are in most cases the major constituents of the sentence.

Finally, the longer durations of the color terms and verbs, and the extra primary stress may function to teach the child how to systematically mark emphasis in his own speech. This is undoubtedly secondary to the direct communicative benefit of these features, but it is something that must be learned at some point since languages differ in the ways in which they express emphasis.

4.2. Age of child listener

As evidenced by the various prosodic differences found between speech to the two year olds and to the five year olds, some of the devices which are commonly used to the first group have disappeared or are greatly diminished in the speech to the second group. A summary of all the characteristics studied and their presence as a function of age appear in Table 11. These observed changes are no doubt due to the linguistic maturation of the child. The age groupings are only meant to be rough indicators of the linguistic abilities of the child and should be considered as such.

The prosodic features that were seen to serve primarily a social function disappeared earlier from the speech of the adults than those serving an analytic one. These include higher fundamental pitch, the use of whispering, and the use of rising final terminals in sentences of imperative form. This is explained by the fact that by the time most children reach the age of four or five, their attention span has improved greatly, eliminating the need for attention getting and attention holding devices on the part of the speaker. Furthermore, by this age probably all of the children have learned the rudimentary rules of conversational exchange and some have already become masters of more sophisticated conversational skills such as verbal manipulation.

Table 11
Presence and absence of some prosodic characteristics in adult speech by function and age of child listener.

Characteristic	Function(s)	Age of Child Listener	
		Two year olds	Five year olds
Higher fundamental pitch	Primarily social	YES	NO
Expanded frequency range	Primarily social	YES	YES
Use of whispering	Primarily social	YES	NO
Rising sentence final pitch terminals	Social/Analytic	YES	NO
Longer durations of verbs	Primarily analytic	YES	NO
of color terms		YES	YES
Use of two primary stresses per sentence unit	Primarily analytic	YES	NO

One feature which does seem to remain is an expanded frequency range, indicating that some instances of high pitch do appear in utterances directed to the five year olds. Not all of the mothers' speech directed to the five year olds exhibited this characteristic. Those that did are in the minority; and these subjects' speech exhibited high pitch peaks in only some utterances. The reasons for these differences between individuals is difficult to determine. My observations indicate that the occasional use of higher pitch to the five year old has little to do with the child's verbal abilities. Instead, it seems to be determined primarily on how interested the child is in the task at hand. When the five year old child is distracted from the task, some mothers will use higher pitch as a device to bring the child back to the task.

By the time the child is five years old, many of the features hypothesized to contribute to the child's analytic endeavor have either disappeared or are greatly reduced in frequency. The emphatic stress on the verbs in the Puzzle task sentences and the use of two primary stresses per sentence are among these. Since the five year old is producing, and therefore presumably fully comprehends sentences of the type in the Puzzle task it is no longer even potentially useful to "cue" the child to the "key" words (the verb) as was necessary earlier. Also, the utterance need not be divided up into such small units as before. The speakers, however, still feel it necessary to modify their speech by prolonging the duration of the color terms in the Puzzle task.

4.3. Speaker interviews

Interviews were conducted with the subjects who participated in the study. One purpose of the interview was to determine if the speakers were aware of the modifications they (and other adults) make in speech to the child listener, and the kinds of differences they noticed. A second purpose was to determine what kinds of verbal means the speakers thought that they used if they specifically wanted to gain the child's attention. It was of interest to see how closely the respondents' answers to the latter question would match their answers to the former and how these related to the findings of the study.

The respondents were all very cooperative and in most cases volunteered a great deal of information about their observations. Many of their answers were well thought out and quite perceptive. One unfortunate drawback in this type of interview, however, was that the respondents did not possess the vocabulary necessary to make precise statements about such matters. As a result, they used terminology which was sometimes vague and ambiguous. For example, many mothers noted that they felt their speech to the child listener was more animated than to the adult, and that their voice was filled with excitement. It is difficult to determine the precise parameters involved when such labels are used (for an excellent discussion of the problem see Crystal 1969). Despite such drawbacks, some statements

made by the respondents are of potential value not only to the interpretation of the experimental results of the study but to the clarification of certain broader issues that are at the moment vague. Data bearing on such matters have been extracted from the interviews and are discussed below.

Most of the subjects noticed differences in their speech as a function of the type of listener (child versus adult) but not all to the same degree. The subjects who addressed the two year olds said they definitely noticed changes in their speech and observed that the differences were quite dramatic. On the other hand, most of the subjects who addressed the five year olds were more tentative about their responses and qualified them with such words as "probably", "perhaps", and "sometimes". Three of these subjects said they doubted that there were any differences at all, but added that when their child was younger they had noticed differences in their speech. The observations of these two groups of subjects are consistent with the findings in this study, that relative to speech directed to the adult, there are many more differences in speech to the two year old than in speech to the five year old.

The kinds of changes that the C1 subjects noted were higher pitch of voice (C102, C103, C105, C107, C110), an expanded range (C103, C105), less volume in voice (C104, C105, C107, C112) and slower speech (C112). Some of these factors were investigated in the present study and found to be empirically valid. Three of the changes noted by the C1 subjects--higher pitch (C203), decreased volume (C202, C212), and slower speech (C206, C206, C208, C210). All the C2 subjects, however, qualified their observations by specifying the particular situation which would trigger these modifications in their speech. One subject (C203), for example, reported that she used a higher pitched voice to her five year old, but only when she was upset about something the child had done and was scolding him. All the C2 subjects who observed differences emphasized that the changes in speech to the five year old were not present at all times. Rather, the changes occurred in particular situations. The most common situations mentioned were: (a) when the child was in a certain state--tired, emotionally upset; (b) when the adult had to restate a request or command, etc., after failure to convey the message adequately; and (c) when the adult was presenting new information to the child, which might be difficult to understand or carry out.

The C1 subjects, on the other hand, admitted using the features they noticed almost always when they were interacting verbally with the two year olds. They reported noticing these characteristics most when they were trying to get the child's attention or when they were in a one-to-one interaction with the child. They believed that they used the features in their speech because it got them results, i.e. it helped make their communication with the child more effective. Several subjects noted that when they failed to use these features, their communication was not understood by the child. As one subject put it,

There are plenty of times I don't stop to think that he's two and I'll just mumble something at him or make some kind of demand on him and don't really think about whether or not he can understand it. And that's when he's most likely not to respond at all. (C106)

All the subjects felt that using the features (higher pitch, etc.) got them results in communicating with the child and admitted using them in their speech, although one half of the speakers said it was probably not good for the child's language development to "talk down to them" in this fashion.

The most commonly mentioned factor triggering the use of these features of the baby talk register was cognizance of the child's presence. This is not a novel observation (Snow, 1972). Certain subjects also mentioned the specific factors that for them determined whether they used the registral features or not in a particular instance. Age and physical size determined the initial level at which an adult would begin an exchange with a child. Several subjects reported occasions on which they spoke to a child of small stature using what they thought to be appropriate features, only to find their speech inappropriate to the child's actual linguistic capability. They claimed making an immediate adjustment in their speech to the verbal abilities of the child. Feedback from the child plays a crucial role in determining the particular registral features selected by the speaker.

Although all the subjects interviewed had something slightly different to say about the devices they used to get a child's attention, they all agreed that changing one's voice so that it would maximally contrast with the ongoing level of speech was the most effective means.

Often I find I have to do something clever to get a young child's attention. And its more effective to do something competely ridiculous or out of the ordinary. Anything that departs from the ordinary or expected. That gets their attention best. (C202)

The devices mentioned ranged from raising the pitch of the voice, talking louder, and wide variation in pitch on one end of speaking softer, slower, and using whispering. The latter approach was associated with a particular style of dealing with the situation.

It's the background I've had. My mother did that to me too. If she really wanted to get my attention she always whispered. I remember that my sister and I always sat up and took notice. I find it works with my children also. (C205)

These persons were in the minority. Most mothers chose the other approach.

It seems then, in general, that foregrounding the speech directed to the child using the devices that produce contrast was what the mothers found through experience to be most effective with young children. This would uphold the interpretation suggested above that some of the prosodic features serve a social function, i.e. to attract and hold the child's attention.

4.4. Implications for further research.

The present study is a preliminary investigation of the prosodic and paralinguistic characteristics of speech directed to children. It only begins to explore the parameters involved. This section presents some opinions on the important issues that need to be explored as subsequent steps in this ongoing investigation. The discussion which follows will be in two parts. The first part suggests further analyses of potential value on the specific material which formed that data for the present study. In it I suggest other variables that could and should be investigated. The second part offers some questions that must be answered as a next step in exploring this particular line of research. Major areas of future research are indicated and a schematic discussion is presented for gathering experimental evidence.

Further analyses. Several factors are not considered in the present study because it is an initial effort in a virtually unexplored area. For one, further investigation of pitch variability is necessary. Frequency range is a limited indicator of variability because it provides no information on the distribution of pitch and is greatly affected by a few extreme cases. Frequency distribution curves for each set of speech samples would provide some evidence on this factor.

Another factor which warrants further study is speech rate. The subjects indicated in the interviews that they spoke slower to the child listener than to the adult. Two measures might suffice here: average number of words per minute for a selection, e.g. the Rainbow passage, and rate of speech during continuous flow of speech.

Finally, more perceptual analyses of the data need to be performed. Specifically, the utterances directed to the child need to be transcribed so that a study could be made of the types of intonation patterns that are used in speech to the child listener, the frequency with which certain patterns are utilized, and the unique features (if any) that are employed. A number of different transcription systems are available. Some of them would undoubtedly be satisfactory for this purpose. There is a fund of information that could be gained from such an analysis, which would also be very useful to the student of the acquisition of non-segmental phonology (Crystal, 1969a).

Further issues. There are at least two questions that, in my opinion, merit further study. The first concerns the identifiability of the baby-talk register. Are adults able to identify a stretch of speech as directed to a child listener or to an adult listener? And, specifically, are the prosodic and paralinguistic features of the

baby-talk register sufficiently salient cues to form the basis for such identification? A study has been begun to answer these questions. It consists of two parts. In the first part subjects listen to short excerpts of speech. Some of these excerpts were taken from speech sequences spoken to an adult listener, and others from sequences spoken to the child. The excerpts are free from all extraneous cues (e.g. the child's voice in background, etc.) except those which are characteristic of the register involved, (e.g. use of special lexical items, repetitions, shorter sentences, etc.). The subjects are asked to label each speech excerpt as "adult listener" or "child listener." In the second part, a different set of subjects listen to repeated presentations of sentences from a list which they have read previously. The sentences are produced by several different speakers. Some of the sentences were recorded when the speaker was addressing the sentence to a child, others when the speaker was addressing an adult. Each pair of sentences is selected so that the segmental aspects of the sentence match one another very closely and that only the prosodic and paralinguistic aspects differ. The sentences are presented to the subjects in random order. The subjects' task is to label each sentence as to whether it was spoken to a child listener or to an adult. The purpose of the latter part of the study is to determine if subjects can identify the age status of the listener to whom speech was directed on the basis of prosodic cues alone.

It is not enough, however, to show that there are prosodic differences between speech directed to a child listener and an adult, and that, furthermore, these differences are salient to the degree that adults can recognize whether excerpts of speech are child directed or adult directed on the basis of these cues. The second question, therefore, concerns the effect, if any, of such prosodic features on the child's attention and performance. One possible procedure for investigating this question would be to place a loud speaker into a large toy animal and have this animal "speak" to the child. The animal could, for instance, give the child directions on completing a task, e.g. putting together a picture puzzle. In one condition, the sentences given as directions would be sentences prerecorded in an adult listener situation. In another condition, the sentences would be those pre-recorded in a child listener situation. The child's response, or the lack of it, would be compared as a function of condition. It is hypothesized that the child would respond more often and more appropriately to directions produced with baby talk characteristics than to the other sentences.

CHAPTER 5: CONCLUSION

5.1. Conclusion

As stated in the introduction, the goal of the present investigation is not only to provide some empirical data on a particular aspect of the linguistic characteristics of speech addressed to young children but also to consider the broader implications of such findings for the theory of language development and linguistic theory more generally. In what follows, the findings of the study will be briefly summarized and discussed in the light of these broader considerations.

5.2. Implications for theory of language development

The first question the present investigation was designed to answer is primarily descriptive: What are the linguistic characteristics of speech to young children? Specifically, what are the prosodic and paralinguistic features which most commonly appear in speech to the young child? Further, how does the use of these features change as the child gets older and thus more sophisticated linguistically? The analysis of speech samples of mothers addressing their two year old or five year old child indicated that speech to the younger children contained (1) higher mean fundamental frequency, (2) a greatly expanded frequency range, (3) numerous instances of rising sentence pitch terminals in declarative and imperative sentences, (4) a high incidence of whispering, (5) increased durations of certain key content words in sentences, and (6) the multiple assignment of primary stress within a sentence unit. Speech to the older children, however, contained only features (2) and (5) and even then the frequency range was less expanded with only one group of words showing longer duration. Although this list of features is preliminary and incomplete it is clear that the adult speaker makes a systematic adjustment in the prosodic and paralinguistic aspects of speech addressed to the young child and that these adjustments vary in relation to the age of the child listener.

With respect to a theory of language development a relevant question which one would pose concerning the characteristics of speech directed to children is whether any of these features of the verbal environment are in some way critical for language learning. Since a definitive experiment to answer this question is not possible on ethical and moral grounds, one can only speculate with greater or lesser assurance, as to whether one feature or another (or more likely some combination of features) may play a significant role. Such speculation would best be directed toward analysis aimed at exploring the function(s) certain features may serve in speech directed to children.

In this investigation the features found in speech to the child listener were classified according to whether they could be interpreted as serving primarily (a) an analytic function, to aid the child in the linguistic analysis of the speech he hears, or (b) a social

function, to capture and maintain the child's attention to the speech directed to him. Features serving a primarily social function were higher fundamental frequency, expanded frequency range and use of whispering. On the other hand, longer durations of content words and assignment of multiple primary stress within a sentence unit were interpreted as serving primarily an analytic function. The characteristic use of rising sentence final pitch terminals was seen to serve both functions.

There is, in my opinion, a stronger argument for considering features serving an analytic function as playing a "critical" role in language acquisition. The social function, after all, can be adequately fulfilled by other means of communication. For example, the use of kinesic cues, i.e., touching, using facial expressions, etc., may be just as effective in capturing and maintaining the child's attention. The features serving primarily an analytic function may not only provide the child with cues necessary for the child's analysis of constituent structure (and the like) but may also be invaluable to the child's learning of certain aspects of the prosodic system of the language he is learning, such as, for example, the system of contrastive stress in English. However, this is not directly true for intonation contour per se. The child has to learn that different utterances (e.g., imperative, declarative, questions) do contrast in adult speech.

Finally, the search for "critical" features for language acquisition may be a misguided venture in itself. What is strongly implied in the question of whether one or another feature is critical is an all-or-none situation. I would suggest that it is more likely that some combination of features is necessary for language acquisition and that each individual feature is only more or less suitable or effective in achieving the goal. Thus, the choice of one subset of effective features may produce a desired result just as effectively as another. As evidenced in the range of individual differences among the adult speakers in the present study, the same effect, e.g., getting the attention of the child, was accomplished in a variety of ways.

5.3. Implications for linguistic theory.

It has been suggested that the study of "simplified registers", of which the baby talk register is one, can contribute to a general linguistic theory by clarifying the notion of simplification in language and the elucidating factors which govern language use (Ferguson 1972). We can now examine how some of the features investigated in the present study could contribute to these endeavors.

Assuming that the hypothesis of universal simplification processes is a productive one, such features of the baby talk register as the prolongation of "key" content words in sentences can be viewed as leading ultimately to the simplification of structure. Longer duration is generally highly correlated with stress placement. Elements that ordinarily carry reduced stress are more often eliminated in the simplification process. For example, articles are often omitted. Thus a sentence type used in the present investigation, e.g., "Push in the blue piece," was consistently produced in such a way that the article "the" was almost totally obscured, while the meaning of the

command was kept intact. Evidence such as this can perhaps be used to explain the omission of the article in languages thought to have undergone some simplification processes, for example, pidgins.

A comparison of the use of prosodic features in the different registers could provide a better understanding of which factors govern the use of particular features. For example, the use of a high pitched voice is a feature peculiar to the baby talk register. None of the other "simplified registers" show a use of this feature. Whatever the reasons for this fact, it seems evident that the presence of a child addressee seems to play a role in the use of this feature. These are only some examples, and perhaps superficial ones at that. As information on "simplified registers" and other registers increases, the contributions of such studies to linguistic theory will become more evident and the contribution to its development should be more widespread.

Finally, the search for "critical" features for language acquisition may be a misguided venture in itself. What is strongly implied in the question of whether one or another feature is critical in an all-or-none situation. I would suggest that it is more likely that some combination of features is necessary for language acquisition and that each individual feature is only more or less suitable or effective in achieving the goal. Thus, the choice of one subset of effective features may produce a desired result just as effectively as another. As evidenced in the range of individual differences among the adult speakers in the present study, the same effect, e.g., getting the attention of the child, was accomplished in a variety of ways.

2.3. Implications for linguistic theory.

It has been suggested that the study of "simplified registers," of which the baby talk register is one, can contribute to a general linguistic theory by identifying the notion of simplification in language and the simplifying factors which govern language use (Vergason 1972). We can now examine how some of the features investigated in the present study could contribute to these endeavors. Assuming that the hypothesis of universal simplification processes is a productive one, such features of the baby talk register as the prolongation of "key" content words in sentences can be viewed as leading ultimately to the simplification of structure. Longer duration is generally highly correlated with stress placement. Elements that ordinarily carry reduced stress are more often eliminated in the simplification process. For example, articles are often omitted. This a sentence type used in the present investigation, e.g., "Push the blue piece," was consistently produced in such a way that the article "the" was almost totally obscured, while the meaning of the

References

- Avram, A. 1967. De la langue gu'on parle aux enfants roumains. In *To Honor Roman Jakobson*, Vol. 1.133-40.
- Austerlitz, R. 1956. Gilyak nursery words. *Word* 12.260-79.
- Bever, T. G., J. A. Fodor, and W. Weksel. 1965. Theoretical notes on the acquisition of syntax: A critique of 'contextual generalization.' *Psychological Review* 72.467-482.
- Bhat, D. N. 1967. Lexical suppletion in baby talk. *Anthropological Linguistics* 9.33-6.
- Blount, B. G. 1972. Parental speech and language acquisition: Some Luo and Samoan examples. *Anthropological Linguistics* 14.119-130.
- Bolinger, D. 1964. Intonation as a universal. In H. Lunt (ed.), *Proceedings of the 9th International Congress of Linguists*. Cambridge, Massachusetts, 1962. The Hague: Mouton, 833-848.
- Broen, P. 1972. The verbal environment of the language learning child. *Monograph of the American Speech and Hearing Association* 17.
- Brown, R. and U. Bellugi. 1964. Three processes in the acquisition of syntax. *Harvard Educational Review* 34.133-151.
- Brown, R., C. B. Cazden, and U. Bellugi. 1969. The child's grammar from I to III. In J. P. Hill (ed.), *Minnesota symposium on child psychology* Vol. 2.28-73. Minneapolis: University of Minnesota Press.
- Byron, J. 1968. Berber nursery language. *Transactions of the Philological Society* 107-61.
- Casagrande, J. B. 1948. Comanche baby language. *International Journal of American Linguistics* 14.11-14.
- Cazden, C. B. 1965. Environmental assistance to the child's acquisition of grammar. Unpublished Ph.D. dissertation, Harvard University.
- Chew, J. J. 1969. The structure of Japanese baby talk. *Journal-Newsletter of the Association of Teachers of Japanese* 6.4-17.
- Chomsky, N. 1959. A review of *Verbal Behavior* by B. F. Skinner. *Language* 35.26-58.
- Chomsky, N. 1967. The formal nature of language. Appendix A to E. H. Lenneberg, *Biological Foundations of Language*. New York: Wiley, 397-442.
- Crawford, J. M. 1970. Cocopa baby talk. *International Journal of American Linguistics* 36.9-13.
- Crystal, D. 1973. Non-segmental phonology in first language acquisition: Methodological problems. *Lingua* 32, No. 1-2.1-45.
- Drach, K. 1969. The language of the parent: A pilot study. *Working Papers No. 14*, Language Behavior Research Laboratory, University of California, Berkeley.
- Drachman, G. 1973. Baby talk in Greek. *OSU WPL* 15.
- Ellis, J. and J. Ure. 1969. Language variety: register. In A. R. Metham, *Encyclopedia of Linguistics information and control*.

- Ervin-Tripp, S. M. 1970. Discourse agreement: How children answer questions. In J. R. Hayes (ed.), *Cognition and the Development of Language*. New York: Wiley, 79-107.
- Fairbanks, G. 1940. *Voice and Articulation Drillbook*. New York: Harper.
- Farwell, C. 1973. The language spoken to children. *Papers and Reports on Child Language Development*. Stanford University 5.31-62.
- Ferguson, C. A. 1956. Arabic baby talk. In Morris Halle, et al., (eds.), *For Roman Jakobson*. The Hague: Mouton 121-128.
- Ferguson, C. A. 1964. Baby talk in six languages. *American Anthropologist* 66.103-114.
- Ferguson, C. A. 1971. Absence of the copula and the notion of simplicity. In D. Hymes (ed.), *Pidginization and Creolization of Language*. Cambridge: Cambridge University Press, 141-150.
- Ferguson, C. A. 1972. Towards a characterization of English foreigner talk. Paper presented at the Third International Congress of Applied Linguistics, Copenhagen, Denmark.
- Ferguson, C. A. 1974. Baby talk as a simplified register. Paper presented at the Conference on Language Input and Acquisition, Boston, Massachusetts.
- Ferguson, C. A., D. B. Peizer, and T. E. Weeks. 1973. Model-and-replica phonological grammar of a child's first words. *Lingua* 31.35-65.
- Fischer, J. L. 1970. Linguistic socialization: Japan and the United States. In R. Hill and R. Konig (eds.), *Families in East and West*. The Hague: Mouton, 107-119.
- Gleason, J. 1973. Code switching in children's language. In T. E. Moore (ed.), *Cognitive development and the acquisition of language*. New York: Academic Press, 159-168.
- Grewel, F. 1959. How do children acquire the use of language? *Phonetica* 3.193-202.
- Halliday, M. A. K., A. McIntosh, and P. D. Strevens. 1964. *The Linguistic Sciences and Language Teaching*. London: Longmans.
- Henry, F. 1948. Discrimination of the duration of a sound. *Journal of Experimental Psychology* 38.734-743.
- Henzl, V. 1973. Linguistic register of foreign language instruction. *Language Learning* 23.207-222.
- Horii, Y. 1972. Fundamental frequency measures estimated from analysis of one sentence voice samples. *Asha* 14.466.
- House, A. S. and G. Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25.105-113.
- Ivić, P. and I. Lehiste. 1969. Prilozi ispitivanju fonestke i fonološke prirode akcenata u savremenom srpskohrvatskom jeziku IV. *Zbornik za filologiju: lingvistiku* 12.115-165.
- Kelkar, A. 1964. Marathi baby talk. *Word* 20.40-54.
- Kobashigawa, B. 1969. Repetitions in a mother's speech to her child. *Working Papers* No. 14, Language Behavior Research Laboratory, University of California
- Kones, R. ms. How to use the pitch extractor with the computer. Phonetics Laboratory, University of California, Berkeley.

- Labov, W. 1970. The study of language in its social context. *Studium Generale* 23.30-87.
- Lehiste, I. 1970. *Suprasegmentals*. Cambridge, Mass.: MIT Press.
- Lehiste, I., and G. E. Peterson. 1961. Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America* 33.419-425.
- Linke, C. E. 1953. A study of pitch characteristics of female voices and their relationship to vocal effectiveness. Unpublished Ph.D. dissertation, University of Iowa.
- Markel, N. N. 1965. The reliability of coding paralanguage: Pitch, Loudness, and tempo. *Journal of Verbal Learning and Verbal Behavior* 4.306-308.
- McNeill, D. 1970. *The acquisition of language: The study of developmental psycholinguistics*. New York: Harper & Row.
- Peterson, G. E. and H. L. Barney. 1952. Control methods used in a study of vowels. *Journal of the Acoustical Society of America* 24.175-184.
- Pfuderer, C. 1969. Some suggestions for a syntactic characterization of baby talk style. *Working Papers* No. 14, Language Behavior Research Laboratory, University of California, Berkeley.
- Phillips, J. R. 1970. Formal characteristics of speech which mothers use to address their young children. Unpublished Ph.D. dissertation, Johns Hopkins University.
- Phillips, J. R. 1973. Syntax and vocabulary of mothers' speech to young children: Age and sex comparisons. *Child Development* 44.183-185.
- Remick, H. L. 1971. The maternal environment of linguistic development. Unpublished Ph.D. dissertation, University of California at Davis.
- Ruhm, H. B., E. O. Mencke, B. Milborn, W. A. Cooper, Jr., and D. E. Rose. 1966. Differential sensitivity to duration of acoustic stimuli. *Journal of Speech and Hearing Research* 9.371-384.
- Ruke-Dravina, V. 1961. Ns. lastenhoitajain kielesta' ("On so-called nursery language"). *Virittaja* 1.85-91.
- Sachs, J., R. Brown, and R. Salerno. 1972. Adults' speech to children. Paper presented at International Symposium on First Language Acquisition. Florence, Italy.
- Skinner, B. V. 1957. *Verbal Behavior*. New York: Appleton-Century-Crofts.
- Slobin, D. I. (ed.), 1967. *A field manual for cross-cultural study of the acquisition of communicative competence*. Berkeley: University of California, ASUC Bookstore.
- Slobin, D. I. 1969. Questions of language development in cross-cultural perspective. *Working Papers* No. 14, Language Behavior Research Laboratory, University of California, Berkeley.
- Snidecor, J. C. 1951. The pitch and duration characteristics of superior female speakers during oral reading. *Journal of Speech and Hearing Disorders* 16.44-52.
- Snow, C. E. 1972. Mothers' speech to children learning language. *Child Development* 43.549-65.

- Snow, C. E. 1974. Mothers' speech research: An overview. Papers presented at the Conference on Language Input and Acquisition. Boston, Massachusetts.
- Stott, L. H. 1935. Time-order errors in the discrimination of short tonal durations. *Journal of Experimental Psychology* 18.741-766.
- Voegelin, C. F., and F. M. Robinett. 1954. 'Mother language' in Hidatsa. *International Journal of American Linguistics* 20.65-70.
- Vorster, J. 1974. Mothers' speech to children: Some methodological considerations. M.A. thesis, Institute for General Linguistics, University of Amsterdam
- Waterson, N. 1971. Child phonology: A prosodic view. *Journal of Linguistics* 7.179-211.
- Willis, D. D. 1974. Participant deixis in English and childish. Paper presented at the Conference on Language Input and Acquisition. Boston, Massachusetts.

Nonverbal Concomitants of Language to Children:
Clues to Meaning*

Olga K. Garnica

Recently there has been increasing interest in the hypothesis that the verbal environment in which language is acquired by the young child may play a significant role in the language learning process itself. Numerous studies have shown that speech directed to the language learning child differs in systematic ways from speech directed to adolescents and adults (Snow, in press; Garnica, 1976, in press). It is suggested that the speech the young child hears, characterized by short and simple sentences, multiple repetitions, prosodic modifications, etc., may be useful to the child in developing an understanding of the precise relationship between meaning and linguistic expression within his language.

To date discussions have focused on strictly the verbal aspects of adult-child communication. The child, however, has a wider range of input available to him. In communicative situations, particularly conversational exchanges, the nonverbal aspects of the face-to-face interaction, e.g. gestures, may be an important source of cues to the speaker's intention. These features of the child's communicative environment have thus far not been studied.

In this paper I analyze one aspect of nonverbal behavior concomitant with verbalizations produced by mothers interacting with their young children. The purpose is to examine the frequency and type of nonverbal cues accompanying verbalizations directed to the young child and to observe how these cues vary with the response of the child as well as with his linguistic sophistication. The paper is part of a larger study designed to specify the informational sources--clues to meaning--available to the child as he proceeds to learn language (Garnica, forthcoming).

1. Method

The data examined in this paper is based on videotaped recordings of nine mother-child dyads. Three of the children were one year olds (19 mos., 18 Mos., 20 mos.), another three were two year olds (26 mos., 31 mos., 32 mos.) and the remaining three were three year olds (38 mos., 40 mos., 42 mos.).¹ The mother-child dyads were videotaped for fifteen minutes through a one-way mirror while they were engaged in an unstructured play situation in a room well-equipped with books, puzzles and various toys. The nature of these interactions can be characterized in Goffman's (1963) terms as focused interaction, "instances of two or more participants in a situation joining each other openly in maintaining a single focus of cognitive and visual

attention--what it sensed as a single mutual activity" (p. 89).

The speech and concomitant nonverbal behavior of both members of the dyad were transcribed. Speech was transcribed into ordinary orthography except for unintelligible utterances produced by the child. These were transcribed phonetically. The nonverbal behaviors were recorded using a system developed for the larger study.² Due to space limitations, the nonverbal portions of the examples are presented in this paper in the form of summary statements.

2. Requests for action

In this paper I discuss the gestures concomitant with one type of illocutionary act--the request for action (Searle 1969). The request for action (henceforth, RA) is one of a more general set of speech acts whereby one person attempts to influence the behavior (or attitude) of another. Specifically in the RA a speaker conveys to an addressee that he (the speaker) wishes the other (the addressee) to perform an act. The RA can be in direct or indirect form. In this paper only direct requests are considered. Indirect requests, including the children's responses, are considered elsewhere (Garnica, ms.). The speech produced by the mother was scanned for direct request forms.

The imperative form utterances were subjected to the simple test established by Garvey (1975), i.e. the imperative is prefaced with a performative tag (I request, I command, I order you to) and judged for appropriateness in the particular context of the utterance. A total of 192 imperative utterances passed this test. The episodes which contained these utterances were then analyzed.

3. Nonverbal actions accompanying requests for action

Requests for action were directed approximately as frequently to all the children in the sample. The frequency and explicitness, however, of the mother's concomitant gestures varied with the age and responsiveness of the child. This can be seen quite clearly by comparing the following two examples from a one year old child and a three year old child:

(i) One year old child.

GENERAL CONTEXT: Child and mother sitting on floor. Child puts small wooden toys dolls into a toy car which is located between him and his mother.

Mother:	<u>VERBAL</u>	<u>NONVERBAL</u>
	Oh boy!	Mother claps hands.
	That's terrific.	
	Very good!	
	Can you push the car?	Mother points to car as she says "car".

Push the car.	Mother touches car and gives it a tap. Car moves slightly. Child watches, then looks to other side of room.
K _____. Watch.	Mother pushes car, turning it around so it is now in front of child and facing away from him. Child watches entire action of sequence.
Vroom. Vroom. Vroom. Vroom. (sound of car engine)	Mother pushes car in four jerky motions and returns it to the starting point.
Push the car.	Mother leans over and looks directly into the child's face.
Vroom. Vroom. Vroom.	No accompanying action by mother.
Push the car?	Mother looks directly at Child Child picks up another toy and examines it.
Push the car.	Mother pushes car in direction away from child and returns it to starting position.
Vroom. Vroom. Vroom. Vroom.	Mother pushes car in four jerky motions and returns it to start position. Child picks up a toy and examines it.
You want to look at that?	Mother looks at toys in child's hand.
Okay (pause)	
Push the car.	Mother pushes car.

(ii) Three year old child.

GENERAL CONTEXT: Mother and child sitting on floor next to one another examining a puzzle which has pieces shaped like various vehicles (bus, ambulance, car, van, etc.). Each vehicle piece can be removed from board. Underneath each piece is a picture of the inside compartment of the vehicle.

VERBAL

NONVERBAL

Mother: Let's look at this puzzle.	Mother places puzzle in front of child.
Child: Puzzle. That's a puzzle. Puzzle.	
Mother: Yes, Where's the bus? Where's but bus at? Can you pick up the bus? Pick up the bus. Pick up the bus.	Child looking at different parts of the puzzle.

Child: Why? Child looks at mother.
 Mother: Let me see that bus.
 I want to see the bus up close.
 Pick up the bus.
 Oh. Look at the people on Mother points to pictured person
 the bus. on the bus.
 Can you pick up the bus?
 Pick up the bus.
 Child: There. Child picks up bus piece and holds
 it up in front of mother's face.
 Mother: Oh! Very good!

These examples are representative of the data. Although both children received a large number of verbal renditions of the RA, the mother of the one year old presents many more nonverbal cues as to what is requested of the child. She quite explicitly models the action requested on six different occasions, i.e. puts the toy car into motion as she repeats the RA. This modelling begins almost immediately after the RA is verbally introduced.

The mother of the three year old refrains from any overt related gestural behaviors until the RA has been repeated many times. Even then, she only points to the object of the RA (rather than picking up the object). Her pointing gesture accompanies the utterance "Look at the people on the bus". She does not model the action requested nor does she indicate the object of the RA in conjunction with the direct request "Pick up the bus." The use of such more subtle or covert nonverbal cues to meaning is the predominant behavior in mother-child pairs involving three year olds. In fact there was only one case of a mother of a three year old performing the requested action simultaneously uttering the RA. The modelling occurred at the end of a long sequence of verbal renditions of the RA, a sequence not unlike the one in the three year old example given above.

There were two major types of adult nonverbal behavior accompanying RAs. One type involves the manipulation of an object (usually one of the toys). The other type only involves the two interactants. The first type is the most common accounting for almost all the RAs noted. For RAs involving objects the adults exhibited two classes of gestures: (1) pointing to the object(s) referred to in the RA, and (2) manipulating the object(s) referred to in a RA, the gesture may occur simultaneously with the production of the word for the object in question or non-simultaneously:

(iii) Two year old

GENERAL CONTEXT: Toy car and small dolls.

VERBAL

Mother: Put the people in
 the car.

NONVERBAL

Mother points to the two toy
 dolls when saying "people" and
 to car when saying "car".

(iv) One year old

GENERAL CONTEXT: Toy car and small dolls.

VERBAL

Mother: Push the car.

NONVERBAL

Mother points to car.

In the case where the adult manipulates an object or objects referred to in the RA, the gesture may also occur simultaneously or nonsimultaneously. Here, however, we see an added dimension. The adult performs the complete action or only some part of the action sequence. Examples of complete and partial actions in the simultaneous category are given below:

(v) One year old - complete action

GENERAL CONTEXT: Toy car and small dolls.

VERBAL

Mother: Push the car.

NONVERBAL

Mother pushes toy car.

(vi) Two year old - partial action

GENERAL CONTEXT: Child searching for pieces of puzzle.

VERBAL

Mother: Tip it [box] upside down.

NONVERBAL

Mother tips box half way and returns it to the upright position.

Maybe we will find the piece.

Child turns box with puzzle pieces upside down and all the pieces fall out.

Examples of complete and partial actions also occurred in the case of the nonsimultaneous category:

(vii) One year old - complete action

GENERAL CONTEXT: Toy doll and toy chair.

VERBAL

Mother: Look.

NONVERBAL

Adult puts doll in chair and takes doll out. Child watches. Adult points to doll.

Mother: Can you do it?
Put the boy in the chair.

(viii) Two year old - partial action

GENERAL CONTEXT: Toy merry-go-round which has two slots for riders.

VERBAL

Mother: Put the people in the merry-go-round and give them a ride.

NONVERBAL

Adult puts merry-go-round in front of child. Positions two toy dolls around perimeter of merry-go-round.

The second major type of adult nonverbal behavior accompanying RAs did not include objects but rather consisted of the adult herself performing the requested action as in (ix) below or of the adult physically manipulating the child into complying with the RA.

(ix) Two years old.

GENERAL CONTEXT: Child searching for missing piece of puzzle.

VERBAL

Mother: Look in the box and see if you can find it.

NONVERBAL

Mother leans over child's shoulder, facing the box referred to.

(x) One year old.

GENERAL CONTEXT: Mother introducing new activity to child.

VERBAL

Mother: Come sit next to me.

NONVERBAL

Mother holds child under his arms and moves child next to her.

The behavior in (ix) appeared consistently in all the episodes with the two and three year olds. The behavior exhibited in (x), i.e., actually physically handling the child, was restricted to the episodes with the one year olds. The other mothers used verbal means ("Come sit here and see what toys I have".) often combined with nonverbal features (e.g. Mother holds up toy with one hand and pats place on floor where she wants child to sit).

4. Concluding remarks

From the analysis of the mother-child interactions, we conclude that the adult adopts certain nonverbal strategies such as modelling, pointing, etc., as an adjustment to the child's limited understanding

of the meaning being conveyed in the utterances directed to him. Further investigation is necessary to determine whether these and other strategies produce a sufficiently rich environment which provides the child with specific information which could be used for testing his/her linguistic hypotheses. The lines of inquiry we are pursuing include a detailed analysis of the sequential patterns established on the adult-child interaction unit which includes RAs (direct and indirect), an examination of other types of speech acts and their concomitant nonverbal aspects, and an analysis of the effect of the type of child response (or lack of it) on the patterning of the mother's nonverbal behavior. With this type of information we can better assess the contribution of aspects of the total communicative environment to the language development of the child.

Footnotes

*Portions of this paper were presented at the Third International Child Language Symposium, London, September 1975. The research is supported by a grant from the Graduate School of The Ohio State University through its Small Grants Program. I would like to thank Louise Charry, Grace Shugar, and Catherine Snow for their comments on an earlier version of this paper.

1. The videotapes were selected for analysis randomly from a set of ten such pairs for each of the three age groups. This material forms the data base of the analysis presented in Garnica (ms., forthcoming).

2. Various parts of this system are still being revised and supplemented. The fully developed notation will appear in Garnica (forthcoming).

References

- Garnica, O. K. 1975. *Some Prosodic Characteristics of Speech to Young Children*. Ph.D. dissertation, Stanford University.
- Garnica, O. K. In press. Some prosodic and paralinguistic features of speech directed to young children. In C. A. Ferguson and C. Snow (eds.), *Talking to Children*. Cambridge University Press.
- Garnica, O. K. Nonverbal concomitants of language input to children II: Indirect requests. ms.
- Garnica, O. K. Forthcoming. *Adult-child discourse: the communicative environment of the language learning child*.
- Garvey, C. Requests and responses in children's speech. *Journal of Child Language* 2.41-63.

Goffman, I. 1963. *Behavior in Public Places*. New York: Freeman Press.
 Searle, J. 1969. *Speech Acts*. Cambridge: Cambridge University Press.
 Snow, C. In press. Mothers' speech research: An overview. In C. A. Ferguson and C. Snow (eds.), *Talking to Children*. Cambridge University Press.

of speech to
 Further investigation
 other
 provides the child with
 for test
 are provided
 established on the adult-child interaction
 (direct and indirect), an examination of other types of speech acts
 and their concomitant conversational aspects, and an analysis of the effect
 of the type of child response (or lack of it) on the patterning of
 the mother's conversational behavior. With this type of information we
 can better assess the contribution of aspects of the total communicative
 environment to the language development of the child.

Footnotes

*Portions of this paper were presented at the Third International
 Child Language Symposium, London, September 1975. The research is
 supported by a grant from the Graduate School of The Ohio State
 University through its Small Grants Program. I would like to thank
 Louise Charry, Grace Shugar, and Catherine Snow for their comments on
 an earlier version of this paper.
 1. The videotapes were selected for analysis randomly from a
 set of ten such pairs for each of the three age groups. This material
 forms the data base of the analysis presented in Gernica (ms., forth-
 coming).
 2. Various parts of this system are still being revised and
 supplemented. The fully developed notation will appear in Gernica
 (forthcoming).

References

Gernica, O. K. 1975. Some prosodic characteristics of speech to
 young children. Ph.D. dissertation, Stanford University.
 Gernica, O. K. In press. Some prosodic and paralinguistic features
 of speech directed to young children. In C. A. Ferguson and C.
 Snow (eds.), *Talking to Children*. Cambridge University Press.
 Gernica, O. K. Prosodic concomitants of language input to children:
 II: Indirect requests. ms.
 Gernica, O. K. (forthcoming). Adult-child discourse: the communicative
 environment of the language learning child.
 Garvey, C. Requests and responses in children's speech. *Journal of*
Child Language 2:41-63.

Phonological Variation in Children's Speech:
The Trade-off Phenomenon*

Olga K. Garnica
Ohio State University

and

Mary Louise Edwards
Casco, Maine

A question of both theoretical and practical importance for the study of phonological development is whether there is a difference in the status of productions rendered spontaneously by the child and those repeated by the child either after an adult model or his own production. The relevant theoretical questions are (1) are all the child's productions mediated by his current phonological system? and (2) what is the role of the adult model in phonological acquisition? The answers to these questions are of practical importance to the investigator studying phonological development, who must decide whether it is more appropriate to include all of the child's identifiable and intelligible productions in phonological analysis, or just some part.

There are conflicting statements in the literature on the status of the child's productions in response to an adult model. Some investigators suggest that the child is capable of phonetically accurate reproductions of utterances under these circumstances, but that such productions are independent of the child's phonological system. Spontaneous productions, those with no immediately preceding model, are, according to this view, less accurate phonologically because they are mediated by the child's primitive phonology. For example, Waterson (1970) reports that her child would often imitate a word with great phonetic accuracy at a given age, but that when this same word was used subsequently it conformed to the child's less sophisticated system. Moskowitz (1970) suggests that all data on child's sound system based on utterances obtained as imitations (productions immediately after the adult model) are probably unreliable and even misleading.

Another view is that spontaneous and imitated productions have the same status and that it is perfectly acceptable to consider both as evidence in phonological studies. For example, Templin (1957) utilized both types of productions in her analysis of children's speech sound articulations. She justified this procedure on the basis of findings from a previous study in which she found similar results in articulation skills, regardless of the types of production used.

Fitzgerald (1966) reports the same percentage of errors for "immediate imitations", (no intervening words between adult model and child's repetition), and "delayed imitations", after interventions ranging from a brief utterance to greater than ten utterances.

The present study considers the spontaneous and repeated productions of several normally developing children. The questions of primary concern are whether the child's production of a word immediately following an adult model significantly differs from his spontaneous production(s) of that word, and, if such differences do exist, what general principles explain them.

1. Method.

Four children participated in this study (two males and two females). They ranged in age from 1;11 to 2;4. They had normal intelligence and normal hearing, and by all indications were progressing normally in their language development.

Each child took part in two testing sessions. In the sessions pictures and various objects familiar to the children were used to elicit single word responses. The child was presented with a picture book, showing common objects and asked to name these objects. The experimenter presented each stimulus item in turn and asked "what's this?" Usually the child gave a response. If the child did not give an immediate response, a second attempt was made to elicit the word in exactly the same manner. Following the child's production of the word, or an unsuccessful attempt to elicit a word, an imitated response was elicited by the experimenter, who said "that's a . Say ". Productions of words were collected from each child ranged from 27 to 35. This includes words that the child introduced himself during the course of the session.

The testing sessions were conducted in a small room with only the child and the experimenter present. The sessions were recorded on a Revox A77 using a Sony ECM-16 electret condenser microphone that was hidden in a vest worn by the child. Each item produced by the child was transcribed in narrow phonetic transcription by two trained transcribers.

For the purposes of our investigation it was necessary to classify the utterances produced by the children. Investigators of child phonology do not generally make more than two-way distinctions between types of utterances. Items are typically classified either as "spontaneous" or "imitated" (cf. Moscowitz, Waterson, and others). What is meant by each of these terms is often left ambiguous. Only Fitzgerald has made a distinction in types of imitation--immediate vs. delayed.

We found the two-way classification inadequate. The addition of a third category (delayed imitation) was only a small improvement. Even with this three way distinction we were unable to classify more than half of the utterances in our corpus. Thus we expanded the classification system to include five categories. These categories are defined as below:

spontaneous (S) - The child's utterance where no adult vowel of the item is present in the previous five minutes.

echoic (E) - The child's production of a word directly after an adult model.

self-repeated (SR) - The child's repetition of his own production of a word when either no adult intervention occurs or the adult requests a repetition without including the model in the request (e.g. "say that again").

delayed-repetition (DR) - The child's production when there is no immediately preceding model by the adult or child, but the word has been produced within a five minute period.

perseveration (P) - The child's production immediately following two or more (SR)s (whether or not the adult attempts to elicit another item).

2. Results.

2.1. Spontaneous versus echoic forms.

There are several alternative hypotheses that can be made about the relationship between spontaneous and echoic productions:

- 1) In all cases there is no difference in form between the echoic and spontaneous productions of an item.
- 2) In some cases the relationship described in (1) holds but in other cases there are differences between the spontaneous and echoic productions of an item.
- 3) In all cases there are differences in the spontaneous vs. echoic productions of an item.

Cases (2) and (3) imply that the adult model has some influence on the child's subsequent productions. We examined our data for all instances of spontaneous and echoic pairs. We calculated the number of items in which the echoic form was the same as the spontaneous form and also the number of items in which the two differed. The following types of changes were considered as differences:

- (a) one segment substituted for another, or a feature change: zebra (S) [diba] (E) [ðiba] (fricative added)
- (b) a segment is inserted:
vest (S) [væt] (E) [væst]
- (c) a segment is deleted:
spoon (S) [spun] (E) [pun]

The results appear in Table 1. Although the finding is based on only a small number of items (especially for subject number 4), it was found that some of the items showed no differences. This supports hypothesis (2) above and suggests that the adult model may have an effect on the child's subsequent echoic production.

Table I

Subject	Number of Spont./Echoic Pairs	Number of Spont./Echoic Pairs Showing Change	Number of Spont./Echoic Pairs Showing No Change
1	28	25 (90%)	3 (10%)
2	25	15 (60%)	10 (40%)
3	24	19 (79%)	5 (21%)
4	11	6 (55%)	5 (45%)

The next question is whether the changes in echoic production are in the direction of the target sound (as defined by the adult's production). All instances of spontaneous/echoic sequences were examined to determine if the target sound was achieved in the echoic form. The results of this analysis appear in Table 2.

Table 2

Subject	Total Number of Segments Undergoing Change	Number of Segments that Hit Target	Number of Segments That Do Not Hit Target	Indeterminate
1	29	13	7	9
2	19	12	5	2
3	28	17	9	2
4	9	5	3	1

An example of a segment reaching the adult target in the child's echoic form is

S1 milk (S) [mo^uk^h]
 (E) [mⁱu^kh] Adult model [milk^h]

In this example the [l] is still absent in the echoic form, but the vowel has reached the target. In other instances the target sound was reached in the spontaneous production but was altered in the echoic production, e.g.

S1 fish (S) [fɪs]
 (E) [fɛʃ] Adult model [fɪʃ]

Here the vowel becomes less like that in the adult model. Some cases were more complex and could not be classified into one of these two categories. Such instances were labelled "indeterminate", e.g.

S1 yellow (S) [lælo]
 (E) [lʌlo] Adult model [jɛlo]

The first vowel is more raised, but it is also more central and thus misses the target.

Thus we find that the echoic forms often do differ from the spontaneous forms and that in many instances the change is in the direction of improvement of the child's inadequate spontaneous renditions of a word. There are, however, notable exceptions to this pattern. Some echoic forms show no change from the spontaneous forms, while others show an apparent deterioration. Some echoic forms show improvement in one part and deterioration in some other part (see fish example above).

2.2. The "trade off" phenomenon.

One interesting phenomenon we observed in our data is best termed "trade off". In "trade off" one sound segment (or feature) in an echoic form more closely matches the appropriate sound in the model while another segment (or feature) in the same utterance diverges from the model. For example:

<u>thumb</u>	(S)	[fʌŋ]	Adult model	[θʌm]
	(E)	[fɒm]		

This example of "trade-off" involves a consonant and a vowel. The consonant reaches its target while the vowel is lowered. Other examples of "trade-off" may involve two or more consonants.

<u>brush</u>	(S)	[bʌs]	Adult model	[brʌʃ]
	(E)	[wʌsʃ]		

Trade-off may also involve more complicated changes.

<u>airplane</u>	(S)	[^h æ p ^h lɛn]	Adult model	[^h ɛr p ^h lɛn]
	(E)	[^h ʌ p ^h wɛn]		

3. Discussion

Thus far we have observed that there are some differences between spontaneous and repeated utterances. It is now necessary to explore the principles governing these differences. We investigated the possibility that the changes between spontaneous and repeated utterances were connected to the position of the segment in the word (e.g. whether it is the initial sounds or final sounds that are altered when the adult model is presented). The evidence did not support this hypothesis. There was no consistent pattern tied to the position of the segment in the word. Position in the utterance played no systematic role in the trade-off cases as well.

Another possibility is that the variability is governed by the child's phonological system. In order to explore this issue, we analyzed the consonant system of each of the four children, based on the utterances produced by each child during three sessions. These

sessions took place within one month of our testing session. They were part of an independent phonological study and were similar to those described above, except that no attempt was made to elicit echoic responses. We formulated substitution rules for each English consonant sound for five positions in the word: initial, final, pre-consonantal, postconsonantal and intervocalic. Each consonant was then classified according to three degrees of stability:

- 1) stable sound - target sound present 75% of the time.
- 2) stable substitute - non-target sound substituted consistently in 75% of the cases.
- 3) variable - two or more non-target substitutes none of which reach the 75% level.

The data indicate that changes taking place between spontaneous and repeated utterances are governed by the stability of the particular segments. Those sounds which are stable in the child's system do not change in echoic and other repeated forms. This is to be expected. This is not necessarily obvious, since, if one suspects that echoic forms exhibit advances over the child's current system, one would expect these sounds to achieve the target sound in the echoic forms, which they do not always do. Those sounds which are termed variable do sometimes change in the echoic forms, but the target sound appears only rarely.

This analysis may help to explain why the notion of "phonologically progressive" echoic forms arose. A comparison of spontaneous and echoic forms which ignores the variability in sounds in the child's productions would show that in some parts of the echoic forms the target sound is achieved. This holds between (S) and (E) for the initial sound in the following example:

<u>juice</u>	(S)	[dʒus]	Adult model [dʒus]
	(E ₁)	[dʒus]	Adult model [dʒus]
	(E ₂)	[dʒus]	Adult model [dʒus]

In this case the initial segment in the (E₁) form is different from the segment in the (S) form and matches the target sound in the adult model. With respect to variation in the child's system, however, we find that word-initial [dʒ] is variable for this child. The variant which occurs in E₁, that is [dʒ], is one of a number of possibilities which is produced by the child in this position. Other variants are [dʒ] and [dʒ]. The appearance of [dʒ] in the echoic form is not necessarily related to the fact that the utterance was produced after an adult model. An examination of the other echoic form (E₂) supports this claim.

This analysis applies to the trade-off cases as well. Trade-off occurs between sounds which are variable. When one sound in the trade-off reached the target, it is simply a manifestation of variability in the child's system and not necessarily an instance of overall improvement or of a production not mediated by the child's system. When another sound in the system seems to exhibit regression, this too is a manifestation of variability. As an example consider

brush (S) [bʌʃ]
 (E) [bʷʌs]

Adult model [brʌʃ]

For this child /b/ is stable in initial position, postconsonantal /r/ is variable and final /s/ is variable. Initial [b] is not affected in the trade-off. The [w] is a variant which alternates with \emptyset (zero) in initial r-clusters. The final [s] is a variant of /ʃ/.

4. Conclusions.

In summary, we examined the spontaneous and repeated utterances of four children and found that there are differences between the spontaneous and echoic forms produced. Interestingly the child's productions immediately following those of the adult do not necessarily approach the model more closely than do spontaneous productions. In certain cases a "trade-off" occurs, i.e. a sound segment in an echoic form more closely approaches the model while another sound segment falls short of the target. This suggests that echoic (and other repeated forms) are not "phonologically progressive", as had been proposed.

Footnote

*This paper was presented at the IVth International Congress of Applied Linguistics, Stuttgart, Germany, August 1975. The material for this study was collected at Stanford University while both authors were associated with the Child Phonology Project (NSF Grant #30962). We would like to thank Jeannie Luckau for assistance in the data sampling and her many useful suggestions and other assistance. An earlier version of this paper was read by Clara Bush, Eve Clark, and Charles Ferguson. We want to thank them for their insightful criticism and suggestions. The final responsibility for the content is, of course, our own.

References

- Fitzgerald, L. K. 1966. Child language: an analysis of imitations by Gã children. Ms.
 Moscowitz, A. 1970. The two-year-old stage in the acquisition of English phonology. *Language* 26.426-41.
 Templin, M. C. 1957. *Certain language skills in children: their development and interrelationships*. Minneapolis, Minnesota: University of Minnesota Press.
 Waterson, N. 1970. Some speech forms of an English child--a phonological study. *Trans. Philol. Soc.* 1-24.

Phonological Differentiation of a Bilingual Child

Roy C. Major

0. Introduction

0.1. Scope and Purpose

Although previously there have been many studies of bilingual children, none deal extensively with the way surface phonetic realizations of the two phonological systems become separate or differentiated. This study will concentrate on the mechanisms by which the systems become differentiated with respect to language and will suggest some reasons why the substitutions take the form they do. The study is based on my analysis of the phonological development of my daughter Sylvia in her acquisition of American English and Brazilian Portuguese from the age one year seven months to two years eight months.

I will deal with the issue of whether a bilingual child has one or two phonological systems. Up to the age of 1;9 (referred to here as Stage I), the same phonological substitutions apply to her production of phonetically identical segments in either language. After the age of 1;9 (referred to as Stages II, III, and IV), very different substitutions affect identical segments of her English and Portuguese. This paper describes Sylvia's phonological differentiation of English and Portuguese and why it takes the form it does.

Many of her English substitutions occur in various English dialects which she has not heard, and many of her Portuguese substitutions occur in various Portuguese dialects (which likewise she had not heard). Her speech seems to reflect what Sapir has called the distinct "genius" of each language--a genius which guides the unique historical "drift" of each language.

0.2. Theoretical Framework.

The theoretical framework employed in this study is David Stampe's natural phonology theory (Stampe 1969, 1973) which postulates a system of phonological processes motivated by innate limitations of the speech capacity. According to this theory, during acquisition the child's mental representation of speech is approximately identical to the adult's. However, due to natural phonological processes reflecting limitations of phonetic capabilities, pronunciation is often very different from the adult. These processes typically fall into opposing pairs: the syntagmatic processes which make sequences of segments easier to pronounce (e.g., the voicing of obstruents in voiced environments eliminates the adjustment of the glottis from voiced to voiceless and back again), and the paradigmatic processes which make individual segments to perceive and often easier to pronounce (e.g., the devoicing of obstruents increases their

perceptual distinctness from vowels and also eases the pronunciation since their oral stricture impedes the air flow which is necessary for voicing). In order to achieve adult pronunciation the child must suppress, limit, or order those processes of the innate system which are not identical to the system of adult speakers of the language.

Some general processes affecting consonants are:

Paradigmatic

- (a) Obstruents become voiceless
[+obs] → [-voi]
- (b) Obstruents become stops
[+obs] → [+stop]
- (c) Anterior coronal consonants become laminal
C
[+anterior] → [+laminal]
[+coronal]

Syntagmatic

- (d) Obstruents become voiced in voiced environments
[+obs] → [+voi]/[+voi]
- (e) Obstruents become continuants after vowels
[+obs] → [+cnt]/V_____
- (f) Sonorants become nasalized in nasal environments
[-obs] → [+nasal]/[+nasal]

Sylvia's speech production is often characterized by much greater effort than an adult, involving relatively greater subglottal pressure as well as greater muscular force in the oral cavity. Process (c) above, which laminalizes anterior coronal consonants, is one possible realization of the tendency to overarticulate or hyperarticulate. If the child uses great force in his articulation of an anterior coronal consonant (such as a dental or alveolar involving the apex or the front laminal region of the tongue), a large portion of the lamina will tend to be forced upward and to make contact with hard palate. This could be because of inertia or because the child is unable to concentrate the force exclusively at the front portion of the tongue. The tendency for children to palatalize dentals noted by Jakobson (1968:78) is quite likely a consequence of hyperarticulation. Sylvia's l is somewhat palatal, but this is because it is a laminally articulated lateral that is simultaneously dental, alveolar, and palatal.

Hyperarticulation is also evident in Sylvia's production of /p/ and /k/. Because of the high pressure in the oral cavity the articulators frequently slip, allowing air to leak out before the closure is willfully released. This results in a very tight bilabial trill in the case of /p/ (although it is definitely not the affricate [pʰ]), and an affricate [kx] for /k/. The non-existence of [+θ] for /t/ is probably due to the fact that the mass of the tongue involved for the production of /t/ is much less than for that of /k/.

The following processes affecting vowels are the result of the work of Patricia Donegan (Donegan, forthcoming, and Miller 1973). The following terminology is employed: color (or chromatic) = palatal and/or labial; palatal = front; labial = round; achromatic = -palatal and -labial; bleach = loss of color; tense = intense color (as in [i] vs. [I]; achromatic vowels, e.g., [ɨ], are therefore always -tense); ! = especially. "!" has a specific meaning in terms of implicational hierarchy. For example, in (g) below !lower means that if the process affects high or mid vowels, it necessarily affects any low vowels, but if the process affects low vowels it does not necessarily affect mid or high vowels. That is, mid, high \supset low, but low $\not\supset$ mid, high.

Paradigmatic

(g) bleaching

Syllabics which are especially low, bi-colored, short, unstressed, and lax lose their palatality and/or their labiality, especially in the environment of vowels of the same color.

$$\left[\begin{array}{l} +\text{syl} \\ !\text{lower} \\ !\text{bicolored} \\ !\text{short} \\ !\text{unstressed} \\ !\text{lax} \end{array} \right] \rightarrow \left[\begin{array}{l} -\text{pal and/or} \\ -\text{lab} \end{array} \right] / \left[\begin{array}{l} ! \\ \text{V} \\ \text{same} \\ \text{color} \end{array} \right]$$

(h) coloring (palatalization)

Non labial syllabics which are especially high and stressed become palatal, especially in the environment of labial vowels.

$$\left[\begin{array}{l} +\text{syl} \\ -\text{lab} \\ !\text{higher} \\ !\text{stressed} \end{array} \right] \rightarrow \left[\begin{array}{l} ! \\ \text{V} \\ +\text{pal} \end{array} \right] / \left[\begin{array}{l} +\text{lab} \end{array} \right]$$

(i) coloring (labialization)

Non palatal syllabics which are especially high and stressed become labial, especially in the environment of palatal vowels.

$$\left[\begin{array}{l} +\text{syl} \\ -\text{pal} \\ !\text{higher} \\ !\text{stressed} \end{array} \right] \rightarrow \left[\begin{array}{l} ! \\ \text{V} \\ +\text{lab} \end{array} \right] / \left[\begin{array}{l} +\text{pal} \end{array} \right]$$

(j) raising

Chromatic syllabics which are especially tense, low, and short are raised, especially in the environment of achromatic or lax vowels.

$$\left[\begin{array}{l} +\text{syl} \\ +\text{chrom} \\ !\text{tense} \\ !\text{lower} \\ !\text{short} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{higher} \\ \text{achromatic} \\ \text{or lax} \end{array} \right] \begin{array}{c} ! \\ \text{V} \end{array}$$

(k) lowering

Syllabics which are especially chromatic, lax, bi-colored, and long are lowered especially in the environment of tense chromatic non-syllabics.

$$\left[\begin{array}{l} +\text{syl} \\ !\text{chromatic} \\ !\text{lax} \\ !\text{bicolored} \\ !\text{long} \end{array} \right] \rightarrow \left[\begin{array}{l} \text{lower} \\ \text{chromatic and} \\ \text{tense} \\ \text{-syl} \end{array} \right] \begin{array}{c} ! \\ \text{V} \end{array}$$

(l) vowel denasalization

Vowels become denasalized.

$$\text{V} \rightarrow [-\text{nas}]$$

As specified above, the paradigmatic processes, if they do not apply regardless of context, may apply dissimilatively. Thus, in (g) above, bleaching is especially likely in vowels adjacent to glides of the same color; e.g., [iy] → [i̥y] ⊃ [uy] → [i̥y].

Syntagmatic

(m) assimilation in terms of color

$$\text{V} \rightarrow [\text{acolor}]/[\text{acolor}]$$

examples: [eu] → [ø] (French)
[au] → [ou] (Portuguese Paulo)

(n) assimilation in terms of height

$$\text{V} \rightarrow [\text{aheight}]/[\text{aheight}]$$

example: [au/ao] → [o] (French au)

(o) vowel nasalization

$$\text{V} \rightarrow [+nas]/[+nas]$$

0.3. Brief History of Sylvia's Language Exposure.

Sylvia was born in Rio de Janeiro, GB, Brazil where she lived until the age of 1;2. During this period her Brazilian mother

and I (an American) both talked to her exclusively in Portuguese. The only exposure to English she received was through overhearing my wife and me converse in English (the language we usually use together). After we moved to the U.S. (when Sylvia was 1;2), I started talking to her mostly in English while her mother has continued in Portuguese. Up to approximately 1;8 Sylvia was cared for at home by us. From 1;8 to 2;1 she was cared for by another Brazilian for about four hours a day. Since 2;1 Sylvia has been attending a day care center run by the Ohio State University. She has no siblings.

Even though Sylvia is exposed to Portuguese every day and understands it well, at present she has largely diminished her use of Portuguese in favor of English. Even when spoken to by monolingual Brazilian children she frequently responds in English.

0.4. Procedures and Notations.

Through my daily interactions with Sylvia which were at least two hours in length, I collected data by transcribing on the spot. Although I did not transcribe every day, since often there was no observable change from one day to the next, when a change did occur I took note of it. In addition to these transcriptions and daily observations I made tape recordings at irregular intervals at all Stages. The total tape time was approximately five hours.

The consonant symbols employed in this paper are those used by the IPA, except where stated otherwise. The vowel symbols used here and their classification is slightly different, especially in terms of height. The tense/lax distinctions apply only to mid and high front and back vowels (not central or low vowels). [ç] indicates any laminal consonant; it is really simultaneously dental, alveolar, and palatal. (See above for discussion on overarticulation).

		Front		Central		Back	
		unr	rnd	unr	rnd	unr	rnd
High	tense	i	y	ɨ	ɛ	u	u
	lax	ɪ					ʊ
Mid	tense	e	ø	ə	ə	ɤ	o
	lax	ɛ		ʌ			ɔ
Low		æ				a	ɒ
		a					ɔ

A segment in "syllable onset" and "syllable offset" refers to the position of the segment in the syllable. Thus, the r's in spray, pray, and ray occur in the syllable onset and will be represented as /.(C)(C)____, while the r's in for, Ford, and Fords occur in the syllable offset and correspondingly will be represented as /____(C)(C).

In some cases a process occurring at one Stage carries over to the following stage. Thus, a process of an earlier stage which

is not mentioned at a later stage has remained the same.

0.5. Appendixes.

Appendix A contains a list of some irregularities, assimilation, consonant and vowel harmony, and metatheses. Appendix B contains a list of Portuguese phonemes with some very brief comments about dialect variation.

1. Stage I: Ages 1;7-1;9.

Sylvia's speech at this stage consists mostly of one and two word utterances, the two word utterances often containing one Portuguese and one English word. The processes are thus far sufficiently general that they apply to phonetically identical segments in both languages.² Thus, there is evidence at this stage that Sylvia does not differentiate phonologically between the two languages.

1.1. Obstruents.

(1) A neutral vowel schwa ([ə]) may be added to a vowel final voiced stop. If the schwa does not occur, the stop is devoiced by (2) below.

$$\text{OPT } \emptyset \rightarrow \begin{matrix} \text{V} \\ \left[\begin{array}{l} -\text{chrom} \\ -\text{low} \end{array} \right] / \left[\begin{array}{l} -\text{cnt} \\ +\text{voi} \end{array} \right] \text{---} \# \end{matrix}$$

[pige] ~ [pik] pig

(2) Obstruents are devoiced in word final position. This process is obligatory if (1) has not applied. In initial and medial positions voiced/voiceless obstruents contrast.

$$[+\text{obs}] \rightarrow [-\text{voi}]/\text{---} \#$$

[pis]

please

[dɔk]

dog

but [papa]

Portuguese papa 'eat, food'

[baba]

baba(dor) 'bib'

(1) bleeds (2) but (1) does not imply that the release of the stop triggers schwa insertion because stops can be devoiced regardless of whether they are released or not. It seems then that schwa insertion prevents devoicing.

(3) A voiced velar stop optionally becomes a continuant intervocally.³

+obs	→ [+lab]	
+cnt		
+dent		
-sib		
		[ba<f]
		[va<f]

+obs	→ [+lab]	
+cnt		
+dent		
-sib		
		bath
		that

(8) Dental fricatives become sibilants when adjacent to high front vowels.

+obs	→ [+sib]/	V
+cnt		+high
+dent		+front
-sib		
		[tʃis]
		[kosil]

+obs	→ [+sib]/	V
+cnt		+high
+dent		+front
-sib		
		teeth
		coffee
		garfu [gaxfu] 'fork'

contrast [gɔfu] and examples under (7)

Sylvia's pronunciation of her name is especially illustrative of (8).

[siɣvʌ] ~ [sizʌ]	Sylvia
(note *[siɣzʌ] and *[sivʌ] don't occur).	

1.2. Sibilants and Shibilants.

Because no conclusive minimal pairs occur, I am reluctant to say [s] and [z] contrast with [ʃ] and [ʒ] respectively. I do, however, suspect they are beginning to become separate since Sylvia's sibilants have a slight whistling sound of a much higher frequency to the ear than do her shibilants. This slight difference continues through Stage IV when even at this late stage the sibilant/shibilant difference is not nearly as great as it is in adult English and Portuguese.⁷

(9) When [s] occurs before a consonant in syllable onset it optionally syllabifies (or it deletes as in (10) below. (In Portuguese there are no tautosyllabic sequences of s plus consonant).

OPT [+sib] → [+syll]/.____C	
[stʌr'i] ~ [tʌr'i]	study
[spun] ~ [pun]	spoon

(10) If s does not syllabify before a consonant in syllable onset it deletes. Thus, (9) bleeds (10). (See (9) above for examples.)

+sib	→ ∅/.____C
-syll	

(11) Voiceless stops are optionally aspirated in either language. (In adult Portuguese voiceless stops are unaspirated).

OPT $\begin{bmatrix} -\text{cnt} \\ -\text{voi} \end{bmatrix} \rightarrow [+ \text{aspirated}]$

[kok] ~ [k^hok]
[papa] ~ [p^hap^ha]
[pun] ~ [p^hun]
[tau] ~ [t^hau]

Coke
papa 'food, eat'
spoon
tchau

(12) Glottal and uvular fricatives are deleted everywhere.⁸

$\begin{bmatrix} \text{C} \\ -\text{ant} \\ -\text{cor} \\ +\text{cnt} \end{bmatrix} \rightarrow \emptyset^9$

[aj]
[pɔtʌ]

hi
porta [pɔtʌ] 'door'

1.3. Liquids¹⁰

(13) Liquids (except flaps) in syllable offset are optionally deleted.

OPT $\begin{bmatrix} \text{C} \\ +\text{liq} \\ -\text{flap} \end{bmatrix} \rightarrow \emptyset / ____ (\text{C})(\text{C}).$

[bɛ]
[mɔ]
[bɔ]

bear
more
ball

(14) Syllable offset laterals (except flaps) are obligatorily delateralized, leaving behind a non-syllabic back rounded vowel.

$\begin{bmatrix} +\text{lat} \\ -\text{flap} \end{bmatrix} \rightarrow [-\text{lat}] / ____ (\text{C})(\text{C}).$

[mɪʏk]¹¹

milk

(15) American English r in syllable offset is optionally dere-troflexed, leaving its labiality.

OPT [+retroflex]¹² $\rightarrow [-\text{retroflex}] / ____ (\text{C})(\text{C}).$

[taɪgɐ] ~ [taɪgɛ]
[bɛr^ɻi] ~ [bɛr^ɪi]

tiger
birdie

Laterals in syllable onset become laminal (as do other anterior, coronal consonants, specified in process (5)). Laterals in this position have a very definite palatal or "clear" quality to them, but they are not as palatal as the adult Portuguese /ʎ/. To the ear they are intermediate between /l/ and /ʎ/.

[lɑit]
[bo.ɫu]¹³

light
bolo 'cake'

(16) American English r in syllable initial position

deretroflexes, leaving a slightly palatal, rounded non-syllabic, which may be represented as [ɣ] or [ø].

[+retroflex] → [-retroflex]/.____
[ɣokɣok] rock rock

(17) Non-nasal flaps become non-flapped laterals.¹⁴

[+flap -nas]	→	[+lat -flap]		
		[pələlə]	¹⁵	<u>arara</u> [ararΛ] 'macaw'
		[nkɛlu]		<u>não quero</u> 'I don't want (it)'
		but [arɔŋɣi]		<u>I found it</u>
		[kɔ·rɪ]		<u>candy</u>

Especially at later stages, the lateral which is substituted for a flap is much more lenis than the lateral which is substituted for /l/ or /ʌ/. This indicates that the processes are ordered (5) then (17). [kɛlu] at a later stage becomes [kɛlu].

(18) Non-nasal flaps become lateral flaps when adjacent to high front vowels.

[+flap -nas]	→	[+lat]	/	[+high +front]	
		[kɪr'i]			<u>kitty</u>

At 1;8, when Sylvia deleted flaps, kitty was pronounced [ki:]. By 1;9 the process had been suppressed.

(19) Liquids in syllable onset consonant clusters are deleted.

[+liquid]	→	ø/.C(C)____	
		[pis]	<u>please</u>
		[basu]	<u>braso</u> [brasu] 'arm'
		[tʃi]	<u>tree</u>

1.4. Nasals

(20) Nasal consonants nasalize adjacent sonorants.

[-obs]	→	[+nas]/[+nas]	
		[nɪ]	<u>knee</u>
		[fɪŋgɛ]	<u>finger</u>
		[mɔŋrɪŋ]	<u>morning</u>

In adult English and Portuguese this process is limited to sonorants preceding the nasal segment. At 1;9 Sylvia has not limited it, but rather it applies to segments preceding or following the nasal. At Stage II nasalization has been limited to a regressive manner.

(21) Nasal consonants optionally delete before obstruents (leaving behind their nasality, as specified by (20)).

C

OPT [+nas] → ∅ / ____ [+obs]
 [fɪŋgə] ~ [fɪŋgə] finger

(22) Short vowels are optionally denasalized.¹⁶

V

OPT [+short] → [-nas]
 [dʌp] ~ [dʌp] ~ [dʌmp] jump
 [mʌki] monkey
 but [fɪŋgə] (never *[fɪŋgə]) finger

1.5. Vowels.

At this stage I shall limit the discussion to some general observations, due to the high frequency of variation. Thus, for process (23) below, variation occurs within the range between the adult representation and the output produced by the process. For example, (23) can bleach [æ] to [a] but also various intermediate variants occur, namely [a] (a laxer front vowel than [æ]) and [a<ɹ] (a low central vowel).

Paradigmatic Processes

(23) Low and mid lax vowels are optionally bleached. (See section 0 for definitions of terms).

V

OPT { [+low] } → [-chrom]
 { [+mid] }
 { [-tense] }

[bæf] ~ [ba<ɹ>f]	<u>bath</u>
[mɔ] ~ [mʌ>] ¹⁷	<u>more</u>
[mɑrsɛlo] ~ [marsalo]	<u>Marcello</u> [mɑrsɛlo]
[ɛuzʌ] ~ [ɑuzʌ] ¹⁸	<u>Elza</u>
[kɑu] ¹⁹	<u>cow</u>

Bleaching does not occur for high and mid tense vowels.

[kek]	<u>cake</u>
[kok]	<u>Coke</u>

Syntagmatic Processes

Vowels tend to assimilate in terms of height and color. These processes have been limited as follows:

(24) Diphthongs with members having the same color monophthongize (assimilate in terms of height), the non-syllabic assimilating to the syllabic.

$$\begin{array}{c} V \\ \left[\begin{array}{c} -\text{syl} \\ \text{acolor} \end{array} \right] \rightarrow [\beta\text{height}] / \left[\begin{array}{c} +\text{syl} \\ \text{acolor} \\ \beta\text{height} \end{array} \right] \text{---} \end{array}$$

[keik] → [kek] cake
 [kouk] → [kok] Coke
 [bɔ̃l] → [bɔ̃ɥ] → [bɔ] ball

At 1;7 two chromatic vowels of the same height were optionally assimilated in terms of frontness.

[milk] → [miɥk] → [miwk] → [mɪk] milk

At 1;8 only [miɥk] occurred.

(25) A neutral vowel plus a labial glide, i.e., [əɥ], assimilate to [ə].²⁰

$$\begin{array}{c} V \qquad V \\ \left[\begin{array}{c} -\text{chrom} \\ -\text{low} \end{array} \right] \left[\begin{array}{c} +\text{lab} \\ -\text{syl} \end{array} \right] \rightarrow [+lab] \emptyset \\ \begin{array}{ccc} 1 & 2 & 1 \quad 2 \\ (14) & (21) & (25) \end{array} \end{array}$$

[pɛnsəl] → [pɛnsəɥ] → [pɛsəɥ] → [pɛsəl] pencil²¹
 (15) (25)
 [taɪgər] → [taɪgəɥ] → [taɪgəl] tiger

Other diphthongs do not assimilate.

[baɪbaɪ] bye bye
 [kaɥ] cow
 [dɔdɔɪ] dodói [dɔdɔɪ] 'it hurts'

2. Stage II: Ages 1;9-2;1.

During this period Sylvia's utterances extend to two and three or more words. Although many are mixed (e.g., "Deixa alone my cocô mine", 'Leave my poo poo alone'), she has apparently become partially aware that there are differences between the two languages. On several occasions I asked her the same question successively in Portuguese and then in English. The responses were negative in both cases but she used the Portuguese negative when responding to the Portuguese question and the English negative when answering the English question.

"Você quer leite?" "Não." [nãũ]
 "Do you want some milk?" "No." [nou]

That the phonological processes are just beginning to become differentiated with respect to language is observable in affricate substitution and monophthongization of vowels. In the following, processes retained unchanged from Stage I are not mentioned.

2.4. Vowels

Diphthongs of the same color in English no longer monophthongize, i.e., (24) has been suppressed.²⁴ For Portuguese utterances, (24) remains optional through Stage IV.

[keik] (not *[kek])	<u>cake</u>
[noy]	<u>no</u>
[boɥ]	<u>ball</u>
[boysʌ] ~ [bosʌ]	<u>bolsa</u> [boysʌ] 'purse'

Process (25), which substitutes [e] for [eɥ], has been suppressed.

[pɛsɐy] (not *[pɛsɛ])	<u>pencil</u>
(note: *[tɔiɣe] is no longer possible for <u>tiger</u> since syllable offset <u>r</u> is now always pronounced).	

3. Stage III: Ages 2;1-2;3.

The utterances at this stage consist mainly of three or more words with very few mixed utterances. ("Quero mamãe" but "I want Mommy"). During this period the phonological processes become even more differentiated with respect to language. Voiceless stops, liquids, as well as the vowels begin to be realized differently in the two languages, indicating the systems are becoming separate. This phonological differentiation appears gradually and can proceed as follows: For segments in which there exists an alternation between the correct (i.e., the adult) and the incorrect pronunciation, the incorrect variant is gradually eliminated, while the correct one persists. This differentiation is readily observable in the case of aspiration of voiceless stops. In Stage I aspiration of voiceless stops was random in both languages, but by Stage III aspiration has been virtually eliminated in Portuguese (as it is in adult speech) while in English it frequently occurs in places where it would be expected in adult pronunciation.

3.1. Obstruents.

Aspiration of voiceless stops, as stated above, has largely been eliminated in Portuguese utterances. Although in English it still remains variable, gradually toward the end of Stage III it appears with increasing frequency in places it would in adult utterances.

[papai]	<u>papai</u> 'Daddy'
[p ^h eip ^h er] ~ [p ^h eip ^h er]	
~ [peip ^h er] ~ [peip ^h er]	<u>paper</u>

Process (6), affrication of anterior coronal stops before high front vowels, has been suppressed in English but not in Portuguese. (In adult Portuguese it is an active process in most dialects).

[t ^h yi] (formerly [t ^h yi])	<u>tree</u>
[dis]	<u>this</u>
but [t ^h itʃiʌ]	<u>titia</u> 'aunty'

Process (4), which substitutes stops for affricates, has been modified. Affricates before non-high front vowels are beginning to appear as palatalized stops but not as full affricates. Thus, chair sounds phonetically between the adult pronunciation of chair and tear. These "partial" affricates persist through Stage IV.

$$(31) \quad [+del \text{ rel}] \rightarrow \begin{bmatrix} +\text{palatalized} \\ -\text{del rel} \end{bmatrix}$$

[tʃr/çr]	<u>chair</u>
([ç] is a palatalized palatal stop)	
[tʃu]	<u>choo choo</u>

Word initial dental fricatives have changed to stops, even when followed by r (eliminating the first optional part of (26) which substitutes labials for dentals). [ð], which has now entered Sylvia's speech, is also realized as a stop. No intervocalic [θ] occurs. Word finally there has been no change. Thus (26) has been revised to:

$$(32) \quad \begin{bmatrix} +\text{obs} \\ +\text{cnt} \\ +\text{dent} \\ -\text{sib} \end{bmatrix} \rightarrow \left\{ \begin{array}{l} [-\text{cnt}]/\text{---} [+son] \\ \{ [+lab] \text{ OPT} \\ [-\text{dent}] \} / \text{---} \end{array} \right\}$$

[tʃi]	<u>tree</u>
[bɑðr]	<u>bother</u>

Process (27), which de-dentalizes labials, persists for [fr] sequences; e.g., Fruit Loops is still often pronounced [fɹutlups].

At this stage we can propose to derive the flap in English from /t/ and /d/ rather than starting with a flap in the underlying representation as suggested in Stage I.²⁵ (In adult Portuguese, flapping of dental stops does not occur. In addition, /t/ and /d/ are distinct from /r/). However, on a few rare occasions I have observed Sylvia flap a Portuguese /d/ but never a /t/.

$$(33) \quad \begin{bmatrix} -\text{cnt} \\ +\text{ant} \\ +\text{cor} \end{bmatrix} \rightarrow [+flap]/[-\text{cns}] \text{---}$$

Condition: the segment must be [+voil] for Portuguese in order to flap.

$$(34) \quad [+flap] \rightarrow [+voil]/\text{---}[+voil]$$

[mæd] ~ [mər]	<u>mad</u>
[æt] ~ [æɾ]	<u>hat</u>
but [aɪgɛlit]	<u>I get it</u>
[bɹigadʌ] ~ [bɹigərʌ]	<u>brigada</u> 'thanks'

Process (17), lateralization of flaps, is now optional in both languages, as well as (28), hardening of flaps to stops.²⁶

[wɔlər] ~ [wɔdər] ~ [wɔrər]	<u>water</u>
[kɛlu] ~ [kɛdu] ~ [kɛru]	<u>quero</u> 'I want'
[dɛmun]	<u>the moon</u>
but [silɛmun] ~ [sirɛmun]	<u>see the moon</u>

The relative rarity of Sylvia's flapping of stops in Portuguese and the relative frequency in English indicate that phonological differentiation is proceeding.

3.2. Liquids

The flap in Portuguese (as discussed above) continues to be realized as [r], [d], or most frequently as [l].

[kɛru] ~ [kɛdu] ~ [kɛlu]	<u>quero</u> 'I want'
--------------------------	-----------------------

In English intervocalic [l], [r], [ɹ]([w]), and [ɪ]([j]) optionally geminate while in Portuguese gemination is limited to [ɹ] and [ɪ]. ([+voc] refers to [l], [r], [ɹ], and [ɪ]).

(35) English	
OPT	V $\begin{bmatrix} +\text{voc} \\ -\text{syl} \end{bmatrix}$ V → 1 2 2 3
	1 2 3

(36) Portuguese	
OPT	V $\begin{bmatrix} +\text{voc} \\ -\text{syl} \\ -\text{lat} \end{bmatrix}$ V → 1 2 2 3
	1 2 3

(Note: the retroflex r does not occur intervocalically in Portuguese and therefore need not be specified)

[maɪjænd] ~ [maɪænd]	<u>my hand</u>
[maɪjæpɛu] ~ [maɪæpɛu]	<u>my apple</u>
[juuwar] ~ [juwar]	<u>you are</u>
[mɔryiɛl] ~ [mɔyiɛl]	<u>Maria</u>

([y] is a front rounded glide which results from deretroflexing r by process (29)).

[taeijadu] ~ [taɛjadu]	<u>tá errado</u> 'it's wrong'
------------------------	-------------------------------

Condition on process (35) (English): If 2 = [+lat] the process is inapplicable if 1 = [-low, +pall]; otherwise the process is obligatory for laterals. After gemination the first segment delateralizes to [ɹ] as specified by (14) and the second becomes laminal as specified by (5). As stated in (36), laterals in Portuguese never geminate.

[dɔɹ.ɫi]	<u>dolly</u>
[æɹ.ɫeɣerər]	<u>alligator</u>
[sæɹ.ɫi]	<u>Sally</u>
but [dʒ.ɫi]	<u>jelly</u>
[sɪ.ɫi]	<u>silly</u>
[fa.ɫʌ]	<u>fala</u> 'talks'

Because the presence of (5) and (14) is demonstrated by such cases as [fɔɹ] fall and [ɫaɪt] light, Sylvia's pronunciation of intervocalic ɫ reflects the difference in syllable structure of the two languages. In Portuguese ɫ occurs in syllable onset whereas in English ɫ may be ambi-syllabic and thus be doubled across a syllable boundary, especially after stress. Given this syllabification, the application of processes (5) and (14) explain the English examples above.²⁷

It is noteworthy that (17), lateralization of flaps, must follow lateral gemination since the laterals which result from flaps are never geminated.

[padi] ~ [pari] ~ [pali] (never *[paɹli])	<u>potty</u>
[dæri] ~ [dæli] (never *[dæɹli])	<u>Daddy</u>
[dɔɹli] ~ [dæɹli] ²⁸	<u>dolly</u>

3.3. Nasals

There has been no noticeable change from Stages I and II.

3.4. Vowels.

At this stage Sylvia's vowels become very different from those in Portuguese. In many of her English utterances the vowels diphthongize by developing (or maintaining) a corresponding glide of the same color and by dissimilating the nucleus. (Cf. dolly above and process (40) in Stage IV). This is especially true of utterances which are produced slowly or lengthened, as in calling from a distance, and utterances with extra heavy stress and/or rapidly rising or falling intonation. Such phenomena do not occur in her Portuguese utterances.

[peɪŋki]	<u>Pinky</u>
[mameɪ]	<u>Mommy</u>
[kukeɪ]	<u>cookie</u>
[spɔɹn]	<u>spoon</u>
[beɪbeɪ] (At 1;4 often [bibɪ])	<u>baby</u>
[dɔɹli] ~ [dæɹli]	<u>dolly</u>

[sej]	see
but [se] (not *[sej])	(vo)sê [se] 'you'

The significance of the different vowel substitutions will be discussed under Stage IV.

4. Stage IV: Ages 2;3-2;8.

There is very little mixing of utterances at this stage except for an occasional vocabulary item which is known only in one language, (e.g., "I want a bala." 'pill'). A continuing differential treatment of superficially identical segments in the two languages indicates increasingly distinct systems for processing English and Portuguese speech. This is most dramatic in the treatment of the vowels.

4.1. Obstruents

Except for a few minor exceptions, aspiration of voiceless stops in English occurs where it does in adult speech and never occurs in Portuguese.

[koka]	<u>caca</u> 'dirt, yucky'
[k ^h at]	<u>cot</u>

Flapping of stops in English continues but lateralization of flaps, process (17), has decreased word internally. Flaps are still frequently lateralized across word boundaries if the first word ends in a flappable stop, but not if it begins the second word. However, if the flappable segment is preceded by a dental plus vowel, instead of lateralizing, the segment will flap, even across word boundaries.²⁹ In Portuguese, flapping of stops no longer occurs, i.e., (33) has been limited to English.

[aiget]	<u>I get it</u>
[aidit]	<u>I did it</u>
[fɛrəʊn]	<u>fell down</u>
[paɪl] (now never *[pali])	<u>potty</u>
[du(ə)ləʒən]	<u>do it again</u>
but [byigadʌ] (not *[byigarʌ])	<u>(o)brigada</u> 'thanks'

Process (12), deletion of non-anterior non-coronal fricatives ([ç] and [h]), has been limited. At 2;3 [ç] and [h] are pronounced utterance initially but not intervocalically and/across word boundaries.

[hænd]	<u>hand</u>
[majjænd] ³⁰	<u>my hand</u>
[t(ə)ejjadu]	tá errado [taexadu]
	'it's wrong'

By 2;4 [x] and [h] are pronounced every; e.g., [mɑihænd] my hand and [tæx/hadu] tá errado.

4.2. Liquids

In English partial suppression of flap lateralization has occurred (as noted above), but Portuguese [r] is still most frequently realized as [l].

[kɛlu] ~ [kɛru] quero 'I want'

(29), which substitutes [u] for liquids in syllable onset clusters, has been suppressed by the end of Stage IV for laterals. English r and the Portuguese flap in these positions are realized as [y] and [l] respectively. The order or appearance of laterals in syllable onset clusters was as follows: At 2;6 [u] was substituted for the liquids in these positions (except r for which [y] was substituted). A few weeks later class was pronounced [kʰlæs]³¹ beside [gʷæs] glass, [gʷɛndʒi] grande 'big', [suip] sleep [fuɑi] fly, [phʷeɪ] play, [bʷæk] black, and [bʷɑsu] braço 'arm'. A few days later glass became [glæs]³² and grande [glɛndʒi]. About a week after this sleep and fly were pronounced [slip] and [flɑi] but [u] continued to substitute for the [l] of play and black, and the [r] of braço. At 2;7 play was [phʷleɪ] beside [bʷæk] black and [bʷɑsu] braço. Finally, a week later black and braço were [blæk] and [blasu].

A stressed syllable in English tends to attract non-syllabic. Since Sylvia's syllable offset r is pronounced [r] while syllable onset r is pronounced [y] or [ɸ], different stress placements result in different pronunciations of r:

	[mɑ.yiə]	<u>Maria</u>
but	[mɛr.i]	<u>Mary</u>
	[ɛr.ən]	<u>Aaron</u>

Extra slow pronunciations are [mɑr.yiɑ] and [mɛr.yi].³³

4.3. Nasals

/n/ and /ɲ/ in Portuguese have now become distinct in Sylvia's utterances, indicating a suppression of (5), the process which laminalizes anterior, coronal consonants. /l/ and /ʎ/, however, are not yet distinct.

	[siyviɲɑ] ~ [siyviɲ]	<u>Sylvinha</u> 'little Sylvia'
	[pɑnu]	<u>pano</u> 'cloth'
but	[molɑdu]	<u>molhado</u> [mo du] 'wet'
	[falɑ]	<u>fala</u> [falɑ] 'talks'

4.4. Vowels.

The fact that Sylvia's vowels continue to be remarkably different in the two languages suggests there is a fundamental difference in the nature of English and Portuguese vowels. This difference can not be ascertained merely by consulting a chart of phonemes--many "identical" vowels (from the phonemic standpoint) are realized quite differently phonetically. Although Sylvia's modifications somewhat exceed or exaggerate those expected or observed in adult dialects (just as she sometimes over-flaps, e.g., my Daddy [mɑjɹæri]), they are changes of the type that would be expected in the respective languages. Although Sylvia's [meɪ] for me is probably not a common pronunciation in American dialects³⁵ her [mi] is widespread in American English. She does not say [mi] and [miɹ], which are pronunciations foreign to English, yet in Portuguese she pronounces aqui [aki] 'here' as [aki] and quite frequently as [aki] and [akiɹ]. The latter are very common pronunciations in Rio de Janeiro, especially in casual speech of young adults (personal observation). To further demonstrate that her changes correspond with expected or observable change in the respective languages, I will give examples of these changes which occur in their respective dialects--dialects which she has not heard. Finally, I will suggest some reasons why the two languages are developing in such diverse fashions.

4.4.1. English Vowels.

(37) Vowels optionally become tense before r, i.e., lax vowels are raised to the next higher tense vowel. This pronunciation is observable in the speech of Baltimore, Md. (Donegan, personal observation).

OPT V → [+tense]/____[+retroflex]

[mor]

[swithɹ/ɔrt]³⁶

[deɪɹ]

more

sweetheart

there

In the last form [ɛ] tenses and diphthongizes to [eɪ], then [r] syllabifies and diphthongizes to [ɛr].

/ai/ and /au/ optionally dissimilate to [ɔɪ] and [əɪ], and then raise to [oɪ] and [eɪ], while /oi/ may raise to [oɪ]. Thus, an achromatic vowel first dissimilates to the opposite color of a following glide (a paradigmatic process) and then it raises (a syntagmatic process of assimilation in term of height). These two processes are stated as follows:

(38) OPT $\left[\begin{array}{l} +\text{syl} \\ -\text{chrom} \end{array} \right] \rightarrow [\alpha\text{color}]/\text{---} \left[\begin{array}{c} \text{V} \\ -\text{syl} \\ +\text{chrom} \\ \beta\text{color} \end{array} \right]$

(where if α = labial, β = palatal and conversely)

$$(39) \text{ OPT } \begin{bmatrix} +\text{syl} \\ +\text{chrom} \\ \alpha\text{color} \end{bmatrix} \rightarrow [+higher] / \text{---} \begin{bmatrix} \text{V} \\ -\text{syl} \\ \beta\text{color} \end{bmatrix}$$

[yoitn&u] right now
 [boi] boy

(where if α = labial, β = palatal and conversely)

Similar changes have been observed in the Outer Banks of North Carolina (Miller, 1973 and Labov et al. 1972).

Sylvia's production of the other vowels maintains the English tense/lax distinction. For example, me is either [mi] or [mei] while bit is either [bit] or [bet] (never *[bet]).

The tense vowels /i/, /e(i)/, /u/, and /o(u)/ which can be represented as outgliding diphthongs, [VY], with segments of identical color, optionally dissimilate in the following manner:

$$(40) \text{ OPT } \begin{bmatrix} +\text{syl} \\ +\text{chrom} \\ \alpha\text{color} \end{bmatrix} \rightarrow [+lower] / \text{---} \begin{bmatrix} \text{V} \\ -\text{syl} \\ \alpha\text{color} \end{bmatrix}$$

[mi] ~ [mei] me
 [bu] ~ [bui] boot
 [de] ~ [dei] day

The degree of lowering is not specified because various degrees have been observed; e.g., me can be [mi], [mi], [mei] and [mei]. An even more dramatic example of lowering occurs when Sylvia imitates a cartoon tiger in a television commercial. She pronounces it's great as [Itsgrrrrait].

The syllabic part of labial vowels may also bleach (in addition to lowering):³⁷

$$(41) \text{ OPT } \begin{bmatrix} +\text{syl} \\ +\text{lab} \end{bmatrix} \rightarrow [-\text{chrom}] / \text{---} \begin{bmatrix} \text{V} \\ -\text{syl} \\ +\text{lab} \end{bmatrix}$$

[bu] ~ [bi] boot

The achromatic may lower also:

$$(42) \text{ OPT } \begin{bmatrix} +\text{syl} \\ -\text{chrom} \end{bmatrix} \rightarrow [+lower] / \text{---} \begin{bmatrix} \text{V} \\ -\text{syl} \\ +\text{lab} \end{bmatrix}$$

[bu] ~ [bi] ~ [be] boot

(40), (41), and (42) may apply in any order since all of the following have been observed.

[bu] ~ [bi] ~ [be] ~ [bou] boot
 [bou] ~ [beu] ~ [bau]³⁸ boat

The lax vowels /ɪ/, /ɛ/, /æ/,³⁹ /ʊ/, and /ɔ/ behave quite differently:

(43) OPT $\left[\begin{array}{c} +\text{syl} \\ -\text{tense} \end{array} \right] \rightarrow [+lower]$ ⁴⁰

[bit] ~ [bɛt] (not *[be(i)t]) bit

(44) Lax vowels optionally develop a neutral vowel offglide.⁴¹ (This is widespread in the U.S. midlands, especially for lax mid and low vowels).

OPT $\emptyset \rightarrow \left[\begin{array}{c} \text{V} \\ -\text{syl} \\ -\text{chrom} \end{array} \right] / \left[\begin{array}{c} +\text{syl} \\ -\text{tense} \end{array} \right] \text{---}$

[biɪt] ~ [biɪət] ~ [bɛɪət] bit

[fuɪt] ~ [fuɪət] foot

[dɔɪg] ~ [dɔɪg] ~ [dɔɪg] dog

[bæɪd] ~ [bæɪd] bad

(45) After an offglide has developed non-high vowels are optionally raised or tensed.

OPT $\left[\begin{array}{c} +\text{syl} \\ +\text{chrom} \\ -\text{high} \\ -\text{tense} \end{array} \right] \rightarrow \left\{ \begin{array}{c} [-\text{low}] \\ [+tense] \end{array} \right\} / \text{---} \left[\begin{array}{c} \text{V} \\ -\text{syl} \\ -\text{chrom} \end{array} \right]$

[mæɪd] ~ [mɛɪd] ~ [mɛɪd] mad

[dɔɪg] ~ [dɔɪg] dog

[hɛɪd] ~ [hɛɪd] head

(44) feed (45) since a neutral glide is necessary for raising, e.g., bad and dog are never *[bɛd] and *[dog]. The condition [+chrom] is necessary because the achromatic low vowel /a/ is never raised, e.g., cot is [kɔt] and [kɔɪt] but never *[kɔɪt] or *[kɔɪt].

4.4.2. Portuguese Vowels.

Unlike the English diphthongs which show the paradigmatic processes of lowering and bleaching, Sylvia's Portuguese diphthongs show a tendency toward the syntagmatic process of assimilation in terms of height, frontness and color.

(46) OPT /ai/, /au/, /ɛi/, and /ɛũ/ tend to slightly assimilate in terms of frontness and color.

[papɔɪ]

papai 'Daddy'

[mãɛɛ]

mamãe 'Mommy'

[pɛɪ] ~ [pɛɪ]

pão 'bread'

[pɔɪlu] ~ [pɔɪlu]

Paulo

The assimilation may proceed even further (especially in fast speech) for vowels followed by labial glides,⁴² probably by successive stages of assimilation in terms of backness and color, e.g., [meu] → [m̥eu] → [m̥eu] or [myu]; or [meu] → [m̥eu] → [m̥eu] → [mou]. Then by (24), monophthongization, [mou] becomes [mo]. The process which produces the final output (excluding (24)) is summarized below:

(47) OPT [+syll] → [+lab]/	$\begin{matrix} \text{V} \\ \left[\begin{array}{c} -\text{syll} \\ +\text{lab} \end{array} \right] \end{matrix}$	
[ɛme/ɛu] ⁴³		é meu [ɛmeu] 'it's mine'
[kɛlumopapai]		quero meu papai 'I want my Daddy'
[kɛlupõũ]		quero pão [põũ] 'I want (some) bread'

Unlike her English uni-colored diphthongs which undergo lowering and bleaching (cf. (40), (41) and (42)), Sylvia's Portuguese uni-colored diphthongs still monophthongize (cf. (24)). That monophthongization is still an active process in her speech can be seen in the examples just given above as well as the one cited below.

[boysA] ~ [bosA] bolsa [boysA] 'purse'

(48) Monophthongs optionally develop a neutral offglide (especially in slow speech or on heavily stressed syllables), but never lower as English lax vowels do. This process is identical to (44) in English except it is not limited to lax vowels.

OPT ∅ →	$\begin{matrix} \text{V} \\ \left[\begin{array}{c} -\text{syll} \\ -\text{chrom} \end{array} \right] \end{matrix} / [+syll] \underline{\quad}$	
[akíə]		<u>aqui</u> [aki] 'here'
[kokôə]		<u>cocô</u> [koko] 'poo poo'

A tendency toward monophthongization and schwa offgliding are active processes in adult Portuguese. The name Roberto is usually pronounced [xobɛxtu] but on too many occasions I heard from my apartment in Rio de Janeiro the mother of a Roberto calling her son from a distance. Phonetically it was [xobɛxtə]

Monophthongization is also observed especially in informal speech in unstressed positions. Não [nõũ] becomes [nũ] when unstressed and Maurício [maurisiu] becomes [morisiu]. A neutral offglide may even occur on phonemic diphthongs through an intermediate process of monophthongization. This is especially noticeable among the "jet set", "hip" or "in" group of teenagers and young adults in Rio de Janeiro (personal observation).

[madéiɾʌ] → [madéɾʌ] → [madéəɾʌ]	<u>madeira</u> 'wood'
[falóu] → [faló] → [falóə]	<u>falou</u> 'you said it man'
[bí]u] → [bíə]u]	<u>bicho</u> 'guy, dude, bug'
[aló] → [alóə]	<u>alô</u> 'hello' (telephone vocabulary)

The different treatment of tense vowels in the two languages poses an interesting question for phonological theory: Why do tense vowels in Portuguese develop inglides whereas in English they develop outglides? What are the conditioning factors?

If one considers all vowels in Portuguese as lax then inglides would be expected. However, /ei/ and /ou/ (as well as /i/, /u/, /e/, and /o/ are very tense sounding to the ear throughout their production, yet they often monophthongize and develop neutral offglides. (Only /ɛ/, /ɑ/, and /ɔ/ sound lax to the ear. Cf. (48)). Tense vowels in English are usually considered phonetically as [V=V̄] or [V̄V̄] (where [V̄] = [+tense], [V̄] = [+lax] and [V̄] = [-syllabic]). From this representation, diphthongization as it occurs in English (with lowering of the first segment and outgliding of the second) would be expected because less tense vowels (i.e., less colorful) tend to lower or bleach more readily than tense vowels. (Cf. process (g) and (k) in section 1). However, the [V̄V̄] would not be expected to develop a neutral schwa offglide (cf. (44), (47), and (48)). However if we consider the Portuguese tense vowels (including the unicolored diphthongs /ei/ and /ou/) as having the more tense element first (i.e., [V̄V̄]) then a schwa offglide would be probable. Phonetically, this would mean the target position is reached much sooner in Portuguese than in English, or that the first element is more peripheral (i.e., has more color) than in English. English tense vowels seem to work into a more peripheral position, while Portuguese tense monophthongs start out more peripheral. This would explain why in English [ou] → [əu] but in Portuguese [ou] → [o] → [oə]. Some acoustic analysis should be done to determine if this is so. On the other hand /ɛi/ and /ɔu/ in adult Portuguese never monophthongize nor subsequently develop offglides; rather they may slightly lower to [ɛ̄ ī] and [ɔ̄ ū]. This observation suggests that these vowels may be represented as [V̄V̄].

Evidence that Portuguese tense vowels have the tensest part in the onset comes from data comparing how vowels affect preceding consonants in the two languages. If the first part of a vowel is tenser (i.e., more peripheral or more color) than the second part, the vowel would be expected to color the preceding consonant more than if the first element were lax or achromatic. The following examples indicate how Portuguese vowels color the preceding consonant in ways unlike English.

Palatalization or affrication of dental stops before high front vowels can be considered consonant coloring. (See Stage I, process (6) which discusses this anticipatory or regressive assimilation).

The process is very prevalent in many dialects of Brazil, /dia/ and /tia/ becoming /dʒiʌ/ and /tʃiʌ/. In English, palatalization and affrication occur but with a following /j/ but not /i/ or /I/. Compare /dɪdʒu/ did you with the impossible *[dʒɪdʒɪdʒɪtʃet] did Edie eat yet.

Palatalization is also observable in other consonants. The name Márcia in slow speech is [mɑksɪjʌ] but in fast speech the following changes are observed: [mɑksɪjʌ] → [mɑksɪʌ] → [mɑksʌ]. By desyllabification, palatalization, and affrication Rogéria becomes: [xɔʒɛrɪjʌ] → [xɔʒɛrɪʌ] → [xɔʒɛʃʌ] → [xɔʒɛʃʌ]. ([rʒ] is a super short flapped affricate).

Labialization of consonants is also evidence that Portuguese consonants anticipate the color of following vowels: sua [suʌ] in a highly reduced form (e.g., in the sentence sua terra é muito bonita 'your land (country) is very pretty') becomes [suʌ] or [sʌ].

In some English dialects the labial vowels are often preceded by an achromatic, e.g., [dɪ̥u] do, [nɪ̥u] new and [lɪ̥u] Lou. These examples suggest that the labial part of the vowel /u/ is reached very late after syllable onset. If this is true then a central vowel would be expected as an intermediate in the transition toward a backward, labial tongue position. If, however, the back labial position is reached very soon after syllable onset, a neutral vowel intermediate would not be expected. In Portuguese, dental plus labial vowels are never pronounced as in the English examples above. Dudu 'Eddie', nu 'naked' and Lu (girl's name) are [d̥u̥], [n̥u̥], and [l̥u̥].

5. Conclusion.

Up to the age of 1;9, the same phonological substitutions applied to Sylvia's production of phonetically identical segments in both English and Portuguese. After the age of 1;9, phonological differentiation developed and with it a distinct set of substitutions became associated with each language.

Sylvia's phonological differentiation has developed in two basic ways. The first involves segments in which free fluctuation exists between correct and incorrect pronunciations. The incorrect variant in each language is gradually eliminated while the correct variant is maintained. The noticeable increase in frequency of voiceless aspirated stops in her English and the corresponding decrease in her Portuguese clearly illustrate this first mechanism (see Stage III).

The second mechanism consists of the appearance of language specific phonological processes previously not observed. The appearance of different processes for each language does not in any way imply that the processes were not present previously, but only that they did not surface, probably because at an early stage other processes took precedence. In English, the suppression of the syntagmatic process of monophthongization allowed the paradigmatic process of diphthongization to operate overtly. (See Stages III and IV, processes (24), (40), and (42)). However, since in Portuguese

monophthongization is never completely suppressed, the diphthongization characteristic of English does not take place.⁴⁴

One other explanation of why English type diphthongization does not take place in her Portuguese is that if it did, many phonological distinctions would be lost. For example if the process which changes /au/ to [æ̥] or [ɛ̥] were allowed to operate in Portuguese as it does in her English, the two words sal [sɑ̥] 'salt' and céu [sɛ̥] 'sky' would become homophones. (See Appendix B for description of Portuguese phonology).

The emergence of language specific vowel processes suggests there is a basic difference in the timing of target positions of the two languages. I have suggested in section 4 that the Portuguese tense vowels reach their target positions much sooner after syllable onset than in English. If this is true, the tensest part in Portuguese (i.e., more peripheral or more colorful part) occurs soon after onset (and would be represented as [V̄V̄]), while in English the tensest part occurs late after onset (and correspondingly would be represented as [VV]).⁴⁵ This fundamental difference predicts the frequent outgliding in English, e.g., /e/ → [e̠], and the ingliding in Portuguese, e.g., /e/ → [e̠]. (Cf. section 0 on vowel processes).

In addition to the basic difference in timing of Portuguese and English vowels, the appearance of language specific phonological processes perhaps suggests that there is a muscular posture characteristic of each language and a whole set of processes is associated with each posture.⁴⁶ There is indirect evidence that is some entity or articulatory posture which operates at a higher than segment level. Sylvia occasionally inserts a Portuguese word into her English utterances but typically the Portuguese word has an "English" pronunciation. However, the same word in her pure Portuguese utterances has a "Portuguese" pronunciation. For example, bala 'pill' can have quite different pronunciations depending on whether the utterance is part of English or Portuguese discourse. "I want a [bɑ̠.lɑ]" but "Quero [bɑ.lɑ]". (See discussion of liquids at Stage III). Her inability to switch postures quickly suggests that there is something operating on a higher than segment level--perhaps it is on the sentence or even the discourse level. Further support for the existence of an "articulatory posture" comes from adult language. On several occasions I have observed that certain Brazilians, when talking Portuguese, typically pronounce an English word with much more Portuguese "accent" than when they are talking pure English.

Additional evidence for a higher posture can be deduced from my own experiences with English and Portuguese. While living in Brazil, I took a trip alone in which I spent a week without talking any English. Upon returning home and after conversing with my wife in English, I experienced great fatigue and physical discomfort in my whole articulatory musculature. In addition, I experience the same type of discomfort when talking Portuguese for the first time after a short period of abstention. My wife has noted a similar phenomenon with her English and Portuguese.

I suggested before that each posture has associated with it a whole set of phonological processes and from these processes one may reasonably predict what the trend is for future sound change in the language. As Sapir has noted "this direction may be inferred, in the main, from the past history of the language" (1921:151), but it may also be inferred from close examination of active processes in adult and child speech. That Sylvia's pronunciation follows the adult trend of the language is evidenced by the fact that many of her processes are those which actually occur in modern Portuguese and English dialects which she has not heard. (See especially Stage IV).

Although this study involves only one child, it does suggest that child phonology can reveal significant things about the adult system. For example, Sylvia's differential treatment of tense/lax vowels suggests this distinction is significant in English. Finally, the study suggests that the two languages have quite different "grooves" or trends for future change. Evidence for this can be found from dialect comparisons as well as from Sylvia's different treatment of the two languages. Perhaps a language's tendency toward a specific type of change is governed by a language specific speech posture and a specific set of phonological processes which is a natural consequence of such a posture.

Appendix A

Irregularities, Assimilation, Consonant and Vowel Harmony, Metatheses

[tuvi]	<u>TV</u>
[ɑʊskyim]	<u>ice cream</u>
[gĩntɑr]	<u>guitar</u>
[stou]	<u>throw</u>
[pɑlɑlʌ]	<u>arara</u> [ɑrɑrʌ] 'macaw'
[ʃigər]	<u>sugar</u>
[dɔjk] ~ [dɔ̃rk]	<u>dark</u>
[pikinindu]	<u>piquininho</u> [pikiniɲu] 'small'
[biki]	<u>bico</u> [biku] 'pacifier'
[bibi]	<u>baby</u>
[pɪnando] ~ [pɪnəno]	<u>piano</u>
[hɪpənəpərəmɛs]	<u>hipopotamus</u>
[lɛmələid]	<u>lemonade</u>
[mʌnənɛ]	<u>banana</u>
[bədədʌ] ⁴⁷	<u>mamadeira</u> 'baby bottle'
[vyɪvər]	<u>river</u>
[wəɪgwən]	<u>wagon</u>
[lɪlɛr] ~ [lɪrr]	<u>little</u>
[dʌmp]	<u>jump</u>
[sɪrrr]	<u>cereal</u>
[pɛrɛr]	<u>purple</u>
[kɛrfɛr]	<u>careful</u>
[styɛstyit] ~ [stɛrmistyit]	<u>Sesame Street</u>
[fʌlɛr]	<u>flower</u>
[nɛkwɪn]	<u>napkin</u>
[pɔrkədɪz]	<u>Portuguese</u> [? or <u>Português</u>]
[ɛmɪnɛl]	<u>animal</u>
[byɛkstɛs]	<u>breakfast</u>
[tɛbɛr]	<u>table</u>
[wɪndɛr]	<u>window</u>
[wɔrm]	<u>arm</u>

Appendix B

Portuguese Phonemes

p	t	k	i	iu	ĩ	u	ui	uu	ũ	(ũĩ)		
b	d	g	e	ei	eu	ẽ	o	oi	ou	õ	õĩ	
f	s	ʃ	χ	ε	ei	eu	ə	əĩ	əũ	ɔ	ɔi	ɔu
	z	ʒ					a	ai	au			
	l	ʎ										
	r											
m	n	ɲ										

Sibilants and Shibilants

Syllable final /s/, /z/, /ʃ/, and /ʒ/ distinctions are neutralized. The segment is voiced when followed by a voiced segment and voiceless if followed by voicelessness. In Rio de Janeiro the segment in this position is [ʃ] or [ʒ] while in São Paulo, Paraná, and generally in the South it is [s] or [z].

mais 'more'

[mɔiʃ]

(Rio)

[mɔis]

(São Paulo, Paraná, and South)

/χ/ (in orthography r initially and finally, rr medially)

In Rio de Janeiro /χ/ is [χ] in all positions (see note 8). In São Paulo and Rio Grande do Sul the segment is realized as the alveolar trill [r]. In Paraná it is generally [χ] in initial and medial positions but in final position [r], [r], or more commonly the retroflex [ʀ] (similar to American English r).

o rato morreu na porta 'the rat died at the door'

[uχɔtumoχeunapɔχtɐ]

(Rio)

[uratumoreunapɔrtɐ]

(São Paulo, Rio Grande do Sul)

[uχɔtumoχeunapɔʀtɐ]

(Paraná)

/l/

Syllable final l is [ɫ] in Rio, Paraná, and São Paulo (in São Paulo it may also be a very velarized l). In Rio Grande do Sul it is [l] or velarized to [ɫ̠]. In Minas Gerais it is often the retroflex [ɫ̠].

falta 'lack'

[fɔɫtɐ]

(Rio, Paraná, São Paulo)

[faltɐ]

(Rio Grande do Sul)

[fɔɫʀtɐ]

(Minas Gerais)

Footnotes

*This is a revised version of my 1976 Ohio State M.A. Thesis, I wish to express my sincere appreciation to my advisor Professor David Stampe for his many helpful suggestions, comments, and encouragement; to Patricia Donegan for her help with the vowels; to Professor Michael L. Geis for his comments; to Professor Olga K. Garnica for her comments and for first encouraging me to undertake this study; and to my wife Elza for her patience. I also wish to give my special thanks to Sylvia for many wonderful hours of pleasure.

1. [ʌ] represents the central vowel of English but which occurs in many American dialects. IPA considers it back.
2. This obviously does not mean physically identical but rather refers to segments which are usually transcribed the same for both languages, e.g., [i] in feet and fita 'tape'.
3. If (1) and (3) both applied then [piyø] for pig would be expected. However, I have not ever heard this pronunciation.
4. This syntagmatic process of anticipatory assimilation is fairly common in the languages of the world. For example Japanese hati 'bee' is [hatʃi]. Portuguese titia is pronounced in some dialects (including my wife's) as [tʃitʃiʌ].
5. Even though Sylvia's adult model of titia is [tʃitʃiʌ], process (6) is necessary since (4) (which stops affricates) would have already applied.
6. An alternate pronunciation of Sylvia occurred several times: [sɪrvʌ]. This suggests l and r are quite closely related. In fact, in a dialect in the state of Minas Gerais syllable final l is pronounced [r]. (See Appendix B). In addition, when I was teaching English in Rio de Janeiro, some students, when attempting to make syllable final L's in English, produced [r], even though their syllable final l in Portuguese was [ɫ] rather than [r].
7. Just how great a difference is necessary to call [s] distinct from [ʃ] depends on the speaker. In Portuguese, sibilants never contrast syllable finally, although they do contrast syllable initially. In Rio de Janeiro syllable final sibilants are pronounced [ʃ] (or [ʒ] if followed by a voiced segment) while in São Paulo they are pronounced [s] (or [z]). However, standard Portuguese used on national radio and television dictates syllable final sibilants be [s] (or [z]). Announcers from Rio try to suppress this "shibilantization" process, sometimes not too successfully. The result is an intermediate between [s] and [ʃ]. But how their speech is interpreted varies. A native of Rio will hear it as an [s], while a speaker from São Paulo will hear it as an [ʃ].
8. Conceivably these deletions consist of two processes: lenition of the uvular or velar fricative to [h] and then deletion. However, since in Sylvia's mother's speech /x/ is usually [h] we may assume that [h] is Sylvia's representation for the segment. In other adults ri 'I laughed' is often [xi] or [χi] while rua 'street' is frequently [huʌ].

9. According to this process, doggie would become [dɔj] (since the input for (12) would be [dɔjɪ] from (3)). However, since I have not observed this pronunciation, this could indicate the ordering of (12) before (3).

10. Many similar process affecting liquids have been described in detail by M. L. Edwards (1970).

11. See discussion of process (24) for an alternative pronunciation of milk.

12. [+retroflex] for r shall refer to the r-coloring of American r, whether it be a true retroflex or a bunched r.

13. In adult Portuguese the l of bolo occurs in syllable onset: [bo.lu].

14. It is questionable whether a flapping rule should have been proposed. In my speech the flap occurs and I have no reason to believe Sylvia would at this age realize (possibly from morphology) that flaps in English are derived from underlying /t/ and /d/. In Portuguese /r/ is distinct from /t/ and /d/.

15. To my knowledge this is the only word in which a consonant was inserted. Possibly this could have resulted from confusion with English parrot.

16. Smith noted a similar process in his son's speech. He stated it as one process, i.e., nasals are deleted before voiceless stops, leaving no nasality. (N. Smith 1973:13).

17. [ʌ̠] is considerably more retracted than [ʌ] of but; it is the sound which would result if [ɔ] were unrounded.

18. Since /ɛ/ elsewhere did not bleach (e.g., bear was never [ba]), perhaps the [y] caused the lowering and retraction. Other evidence is that meu 'mine' is pronounced [mɛy] by Sylvia. Also see note 19 concerning cow.

19. This example is cited here since the adult pronunciation of cow is either [kɔy] or [kəy].

20. The sequence schwa plus palatal glide, [əj] does not occur in my dialect of English. In Portuguese, the nasal /õ̃/ occurs as [õ̃j] in Sylvia's speech, e.g., [mã̃õ̃j] mamãe 'Mommy'.

21. An alternative analysis would be to start with [pɛsil]. Then by delateralization, [əl] would be the result without having to propose process (25). However, since at Stage II [pɛsɛj] does occur, (25) seems necessary.

Another alternative would be to start with [pɛsɪl] but this is rejected since [ɪu] → [ɪ̃] and not [əl], e.g., milk was [mɪk] but never *[mɛk].

22. Presumably the labial articulation of r is responsible for the labialization of [θ]. Although there were no examples like thwart ([θw]) to check this, David Stampe (personal communication) reports a similar change in his son's speech which affected [sw], e.g., [fwɪm] swim. Also see Stage III where [θ] becomes [d] before [r].

23. The substitution of [ɣ] for [r] after obstruents is fairly common in the speech of residents of the state of Minas Gerais, Brazil (personal observation). In addition, I have observed several

native speakers of Rio de Janeiro substitute [y] for [r] everywhere. Thus carro [kaxu] 'car' and caro [kafu] 'expensive' were distinguished on the basis of voicing, i.e., [kaxu] and kaxu] respectively.

24. An alternative in English would be to start with a monophthong as the underlying representation and state a diphthongizing process.

25. See note 14 concerning Stage I.

26. Although flaps in English are frequently realized as lateral, the flap which results from Portuguese /d/ is never lateralized by Sylvia. But since she flaps stops so rarely in Portuguese, I hesitate to propose the ordering of (17) before (33). For her English utterances, an unconstrained order of (17), lateralization of flaps, (32), stopping of dental fricatives, and (33), flapping, would explain all of Sylvia's pronunciations of [ð] in don't bother me, namely [d], [r], and [l].

27. See Stage IV for significance of r on syllable structure.

28. See Stage IV for explanation of [ou] → [æu].

29. Did it [dɪrɪt] possibly indicates an instance of consonant harmony since consonant harmony has been observed for other consonants, namely nasals and laterals, e.g., [lɛmðleɪd] lemonade. See Appendix A for further examples.

30. The glide gemination is a result of process (35).

31. Also at this time like this (formerly [lɔɪkɪs]) became [lɔɪklɪs]. This is from a progressive assimilation of the l of like rather than the sequence [ð] → [d] → [r] → [l] → [ɹ] as I had originally thought. Evidence for this is that on several occasions in echoic productions Sylvia said [lɔɪklɪs] like this but [pækɪs] pack this.

32. Possibly the more lenis articulation of /g/ vs. /k/ explains this. A similar reason would explain why [pl] appeared before [bl].

33. Again resonant gemination is a result of (35).

34. The process [ɹ] → [j] is common in Munda languages in syllable offset positions. (Stampe, personal communication). [j] is also an acceptable pronunciation of [ɹ] in adult Portuguese.

35. Patricia Donegan and David Stampe report that they have heard this pronunciation in lower-class London kid speech (personal communication).

Paradigmatic processes are also readily observed in singing. In Ray Charles' recording of "Working for the City" (c. 1975) his pronunciation of city changes from [sɪtɪ] to [sɪtɛ/ɛɪ] and even to [sɪtəɪ].

36. Consider rounding of back vowels as a form of tensing.

37. Bleaching does not occur in the syllabic segment for palatal vowels, e.g., me and day never become *[mɛɪ] and *[dɛɪ].

38. Since the lower vowels are more likely to bleach (cf. (23)) this may explain the non-occurrence of [b ut].

39. If /ə/ is treated as [+tense], then (23), bleaching of low or mid lax vowels, would explain [æ] → [ɑ].

40. Since the specification [-tense] does not change, this process implies the vowel is lowered to the next lax vowel.

41. The height specification is not given since it depends on the syllabic which preceeds.
42. Pão [põ] becomes [põ] but mãe never becomes *[mẽ].
43. [ɛ̃] → [ɛ̄] can be viewed as assimilatory retraction since [ɛ̄], although being lower, is more central than [ɛ].
44. In fact monophthongization operates in adult Portuguese (See section 4).
45. Cf. section 4, pp. for further discussion. Also see Lehiste and Peterson's acoustical comparisons of English tense and lax vowels (Lehiste and Peterson 1961).
46. Cf. Drachman (1973) for a discussion of "basis of articulation".
47. [õ] is a "cold" m or the nasal one makes when his nasal passages are blocked.

References

- Burling, R. 1973. Language Development of a Garo and English-speaking child. In C. A. Ferguson and D. I. Slobin (eds.) *Studies of Child Language Development*. New York: Holt, Rinehart and Winston.
- Câmara, J. M. 1970. *Estrutura de Língua Portuguesa*. Petrópolis, R.J. Brazil. Editora Vozes Ltd.
- Chao, Y. R. 1973. The Cantonian idiolect: An analysis of the Chinese spoken by a twenty-eight-month-old child. In C. A. Ferguson and D. I. Slobin (eds.) *Studies of Child Language Development*. New York: Holt, Rinehart and Winston.
- Chomsky, N., and M. Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.
- Delattre, P. 1951. *Principes de Phonétique Française*. Middlebury College, Vermont.
- _____. 1953. Les Modes Phonétiques du Français. *The French Review* XXVII.1.59-63.
- Donegan, P. forthcoming. *The Natural Phonology of Vowels*. Ph.D. Dissertation, Ohio State University, Columbus.
- Drachman, G. 1973. Phonology and the basis of articulation. *OSU WPL* 15.
- Edwards, M. L. 1970. *The Acquisition of Liquids*. M.A. thesis, Ohio State University, Columbus.
- Garnica, O. K. 1973. The development of phonemic speech perception. In T. Moore (ed.) *Cognitive Development and the Acquisition of Language*.
- Imedadze, N. A. 1967. On the psychological nature of child speech formation under condition of exposure to two languages. *International Journal of Psychology* 2.129-132.
- Jakobson, R. 1968. *Child Language, Aphasia, and Phonological Universals*.
- Lehiste, I., and G. Peterson. 1961. Transition, Glides, and diphthongs. *JASA* 33.268-277.

- Leopold, W. F. 1947. *Speech Development of a Bilingual Child*.
Vol. 2. Evanston: Northwestern University Press.
- Labov, W., M. Yaeger, and R. Steiner. 1972. *A Qualitative Study of
Sound Change in Progress*. NSF Report. University of Pennsylvania.
- Malmberg, B. 1963. *Phonetics*. New York.
- Miller, Patricia Donegan. 1973. Bleaching and coloring. *Papers
from the 9th Regional Meeting, Chicago Linguistic Society*
386-397.
- _____. Southern Discomfort. Paper presented at the LSA
meeting, December 1974.
- Sapir, E. 1921. *Language*. New York: Harcourt, Brace & World, Inc.
- Smith, N. 1973. *The Acquisition of Phonology*. Cambridge: Cambridge
University Press.
- Stampe, D. 1969. The acquisition of phonetic representation.
Papers from the 5th Regional Meeting, Chicago Linguistic Society
443-454.
- _____. 1973. *Dissertation on Natural Phonology*. Unpublished
Ph.D. Dissertation, University of Chicago.
- Swain, M. and M. Wesche. 1973. Linguistic interaction: A case study
of a bilingual child. *Working Papers in Bilingualism* 1.10-34.
- Sweet, H. 1892. *A Primer of Phonetics*. Oxford.
- Weir, R. 1970. *Language in the Crib*. The Hague: Mouton.
- Weinreich, U. 1968. *Languages in Contact*. The Hague: Mouton.

A Sociolinguistic Glance at the
Great Vowel Shift of English

John Perkins

The Great Vowel Shift of English has probably been the focal point for more controversy and speculation among historical phonologists than any other phenomenon in the history of the English language. It is universally agreed that between the stage of the language called "Middle English" and the stage referred to as "Modern English" the following changes occurred in the vowel system of the language:¹

- (a) Late Middle English (LME) \bar{i} (Phonetically [i:] was diphthongized by the Fifteenth Century and its nucleus was lowered subsequently to [a], giving modern [aɪ] by the Seventeenth Century.
- (b) LME \bar{u} (phonetically [u:]) was also diphthongized by the Fifteenth Century and its nucleus was eventually lowered to [a], giving modern [aʊ] by the Seventeenth Century.²
- (c) Subsequent to the diphthongization of LME \bar{i} , LME \bar{e} (phonetically [e:]) was raised to [i:] in the Fifteenth Century.
- (d) Next, LME $\bar{\epsilon}$ (phonetically [ɛ:]) was raised to [e:], or tensed, in the Late Fifteenth Century.³
- (e) Subsequent to the diphthongization of LME \bar{u} , LME \bar{o} (phonetically [o:]) was raised to [u:] in the Early Sixteenth Century.
- (f) LME $\bar{\omega}$, (phonetically [ɔ:]) was raised to [o:] in the Sixteenth Century, which was diphthongized to [ou] in the Seventeenth Century.
- (g) Later on, in the Late Seventeenth Century, the new [e:] from LME [ɛ:] was raised even more until it was merged with the [i:] from LME [e:].
- (h) At about the same time, Early Modern English [a:]--the result of a lengthening of LME [a] in open syllables--and [aɪ]--which resulted from a merger of LME \bar{a} and \bar{e} --were each changed in such ways that they were eventually merged as modern [eɪ].
- (i) Finally, also in the Seventeenth Century, LME [aʊ] was monophthongized to [ɔ:].

Diagrammatically, the Vowel Shift can be portrayed as a rotation of the long (tense) vowels and diphthongs in articulatory space:

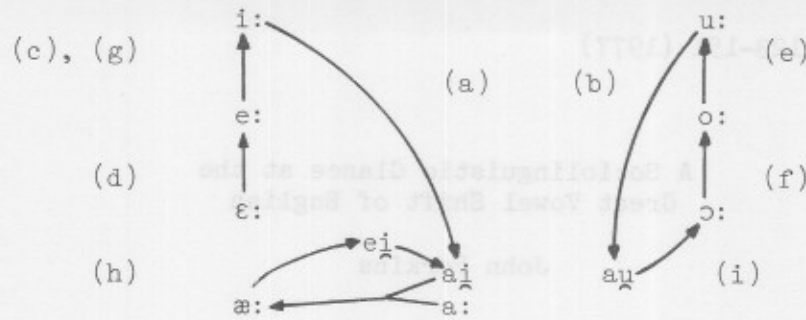


Figure 1.

Philological evidence regarding this series of alterations in the pronunciations of the English vowels was adduced by scholars in the Nineteenth Century, most notably by Ellis (1874); however, it was Otto Jespersen (1909), the originator of the designation "Great Vowel Shift," who first presented an account that considered these changes to be a unified phonological phenomenon. Zachrisson (1913), Wyld (1927, 1936), Kökeritz (1953) and Dobson (1957) have inferred a great deal about the phonetic details of the Vowel Shift from the phonetic accounts of Sixteenth and Seventeenth Century orthoepists and language teachers, as well as Early Modern English misspellings, rhymes and puns.⁴ Linguists have been far from unanimous in their interpretation of the evidence. Their disagreement has centered in particular around three points: the chronology of the Vowel Shift, the phonological motivation of the changes, and the phonetic details of the diphthongization of LME [i:] and [u:].

Wyld (1936:144-145) has argued for a much earlier dating of the Vowel Shift than the generally accepted dating which is presented above. His interpretation of Late Middle English variations in spelling leads him to place the beginnings of the Shift in the late Thirteen-Hundreds.

Most authorities agree that the Vowel Shift began with the diphthongization of LME [i:] and [u:] to [ij] and [uj]. Jespersen felt that this change created "gaps" in the phonological system of long vowels which initiated a "drag chain" that pulled the lower monophthongs up to restore the system. Martinet (1955) adopted Jespersen's drag chain hypothesis and explained it as being motivated by an inherent phonological tendency toward the optimal utilization of phonetic space to maintain phonemic contrasts. Luick, on the other hand, regarded the raising of LME [e:] and [o:] as the initial change and hypothesized that this raising initiated a "push chain" effect that crowded LME [i:] and [u:] out of their positions in the vowel pattern. Trnka (1959) has explained the Vowel Shift as being motivated by the disappearance of the phonological correlation of length and a subsequent reorganization of the phonetic vowel system to restore the correlation. More recently, Chomsky and Halle (1968), Jones (1972) and other generative phonologists have argued that the diphthongization of LME [i:] and [u:] and the exchange of the nuclei of the resulting diphthongs [ij] and [uj] with [e:] and [o:] should be understood as brought about by the addition of phonological rules to the grammar of

English. Strangely, generative phonologists have haggled over the adequate formalization of these putative rules to the neglect of the fact that formal constructs such as "rule addition" are not in themselves explanations of sound change. That is, the fact that the rule content of the phonological component of a grammar must be described differently for two different historical stages of a language is merely a reflection of the fact that there is an intervening phonological change.

Stampe (1972 and personal communication), like Trnka (1959), also hypothesizes that the Vowel Shift was a response to the incipient loss of phonemic vowel length in Late Middle English. As he and Lass (1974) have noted, the length contrast was gradually eroded during the history of English by a series of shortenings and lengthenings which Lass has labelled the "Great English Length Conspiracy." These changes are summarized in Table 1.

	V	V
	[-long] > [+long]	[+long] > [-long]
West Germanic		
Final	/ _ #	
Lengthening:	($\bar{p}\bar{u}$ > $\dot{p}\bar{u}$; $tw\bar{a}$ > $tw\bar{a}$)	
<hr/>		
Old English		
Quantity	/ _ C C	/ _ CC
Adjustment:	[resonant][obstruent]	$\left\{ \begin{array}{l} C \\ VC_1VC_0\# \end{array} \right.$
	(cild > child; fīndan > findan)	(gōdspell > gödspell (*ǣnlefan > ěnlefan)
<hr/>		
Middle English		
Quantity		
Adjustment:	/ C ₀ CVC ₀ #	/ _ C
("Open-Syllable Lengthening")	(ēten > ěten)	$\left\{ \begin{array}{l} C \\ VC_1VC_0\# \end{array} \right.$
		(cēpte > kēpt; mētte > mētte) (hāligdæg > hālidai)

Table 1. The Great English Length Conspiracy (after Lass 1974).

According to Stampe, the motivation of this phonological "conspiracy" was the tendency in a stress-timed language such as English toward isochrony, that is, the equalization of the amount of time allotted between stressed syllables. The outcome of the "Length Conspiracy" was the neutralization of the length opposition in all phonological environments save one, namely, in monosyllabic lexical items ending in a single consonant. Stampe maintains that the "tensing" of the remaining long vowels and the concomitant "laxing" of the short vowels had the effect of sharpening and preserving the perceptibility of the contrast between monophthongal vowel phonemes. Furthermore, Stampe (1972) proposes that the Great Vowel Shift can be understood as the product of innate, natural phonological tendencies which are detailed in Miller (1972).

Miller (1972) has theorized that the tensing of long vowels, especially higher ones, is one of a number of innate phonological processes which operate in the central nervous system to adjust phonological features so that paradigmatic contrasts are more clearly defined in speech. She considers diphthongization of tense vowels to be another "natural process" which increases the "coloration," i.e., contrastive properties, of vocalic phonemes. The lowering of diphthongal nuclei is viewed by Miller as a third process which increases the "sonority," i.e., audibility, of the nuclei and, hence, of the phonemic units. A fourth process is the raising of tense non-high vowels, also viewed as a means of increasing their phonological properties.⁵ Taken together in the order mentioned, these processes can be used to explain the changes involved in the English Vowel Shift.

Although the diphthongization of LME [i:] and [u:] to [iɪ] and [uɯ] is generally accepted to have been the first step in the Vowel Shift, several different hypotheses have been proposed regarding the probable intervening pronunciations of the diphthongs as they were changed to [aɪ] and [aʊ]. The different proposals that have been advanced are distinguished in Figure 2.

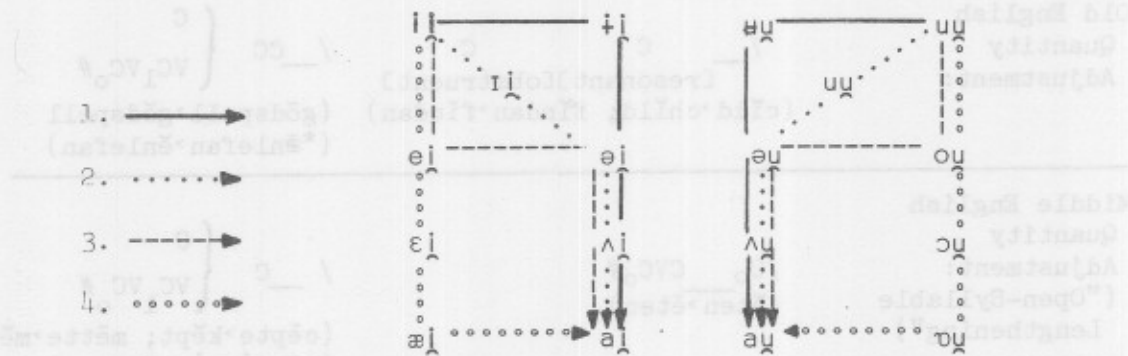


Figure 2.

Wolfe's exhaustive reanalysis of the testimony of the earliest writers on English pronunciation (1972) leads her to the conclusion that only hypothesis 3 is supported by the evidence. Labov, Yaeger and Steiner (1972), however, maintain that their empirical studies of contemporary vowel shifts in progress lend feasibility only to hypothesis 4.

Despite the speculations concerning the phonological factors motivating the Great Vowel Shift, little attention has been focused upon what Labov (1972) has called the "actuation problem," with respect to these changes. In other words, little explanation has been offered as to why the initial changes of [i:] and [u:] occurred when they occurred, and not before. Even if Stampe's theory regarding the causation of the Vowel Shift by innate phonological tendencies is assumed, there remains the problem of explaining why the outputs of certain natural processes were accepted by English-speaking

communities at one particular stage in the history of the language, rather than being repressed as they formerly had been. As Labov (1972) has pointed out, the success of a phonetic change can probably be genuinely understood only if the transition from the earlier pronunciation to the later pronunciation is examined with reference to the social environment within which the change is embedded. Labov (1963, 1966) has produced convincing evidence that alternate phonetic realizations of phonemes play a significant role in marking the social identity of the members of a language community, and that phonological elements which perform this sort of sociolinguistic function are especially amenable to phonetic change. He has shown (Labov 1972) that the members of a community who aspire to higher ascribed status within a group (in whatever terms that status may be defined) will tend to adopt those phonetic variants which are identified as characteristic markers of higher-status speakers. Labov's studies further provide carefully-gathered empirical data which substantiate the long-accepted notion that adoption of pronunciations from one dialect into another characteristically results in the hyperextension of the target variants into environments where they do not actually occur in the speech of the emulated group. In fact, Labov (1966) has gone a step further in suggesting that the generalizing effect of hyperextension is one of the major factors which accelerates the transition from one pronunciation to another.

There is every reason to believe that in the past, phonetic variants functioned as sociolinguistic variables just as they do today, at least in societies with comparable stratificational patterns. The major obstacle to appealing to sociolinguistic factors in attempting to explain past sound changes is the fact that both the phonological variation and the social variation which need to be considered are largely or completely inaccessible to investigation. Where it is, however, possible to reconstruct the social and phonological details of a past age in which a sound change is known to have occurred, it seems clear that significant correlations may be established. The establishment of such correlations is bound to increase our understanding of the actuation and transition problems vis-à-vis particular changes.

In what follows, I will attempt to present the significant details of the sociolinguistic environment within which the beginnings of the Great English Vowel Shift were embedded. I will also try to demonstrate how an understanding of the sociolinguistic situation surrounding historic changes enables us to understand the transition from earlier to later pronunciations.

The Sociolinguistic Situation in Late Medieval England.

From the time of the introduction of feudalism by the Normans in 1066 until the late Thirteenth Century, English society was a rigidly stratified estate system. Authorities on the Norman Occupation of England agree that William the Conqueror replaced the English nobility and clergy virtually in toto with his own French-speaking allies and kinsmen. Loyn (1967) notes that those native

English landowners who survived managed to sustain themselves only by intermarrying with the Normans. This meant, of course, that essentially all the agricultural land, and hence all wealth was concentrated in the hands of French-speaking individuals. The "third-estate," that is, the English and Celtic-speaking farmers and craftsmen who worked for the nobility and clergy in return for protection and sustenance, had no hope for upward social mobility in the first two centuries of the Norman ascendancy. The social barriers to mobility between the classes were, of course, strengthened by the language barrier.

The direct testimony of Medieval English writers indicates that the linguistic divisions of English society continued to conform to the lines of social stratification for many years. Robert of Gloucester, writing in 1300, observes, regarding the Conquerors:

"...The Normans could speak only their own language then
 [i.e., at the time of the Conquest]
 And spoke French as they did at home and also taught [it to]
 their children;
 So that high men of this land that come from their blood
 All keep to that same language that they brought from home.
 For unless a man knows French, little is thought of him.
 But low men keep to English and to their own speech yet..."
 (Mossé 1968, my translation)

Even though the flow of literature in English never completely ceased, the fact that the written language became much more regionally diversified after the Conquest than it had previously been indicates that there was no prestigious form of the language upon which a literary standard could be based. French spelling conventions replaced native spelling conventions and the native "insular miniscule" script was gradually abandoned in favor of the continental "Carolingian" script. These facts, considered together with the predominance of Norman-French literature, argue that the primary colloquial language in which scribes were accustomed to writing was French, rather than English.

As Jones (1972) and Baugh (1957) point out, however, English society must have become increasingly bilingual as time went on.⁷ The following observation is contained in John of Trevisa's English translation (ca. 1368) of Higden's Polychronicon, written about 1327:

"...Children in school, contrary to the custom and manner of all other nations, are compelled to abandon their own language and to construe their lessons and their things in French, and have since the Normans first came to England. Also the children of gentlemen are taught to speak French from the time that they are rocked in their cradles and know how to talk and play with a child's brooch; and rural men want to liken themselves to gentlemen, and apply themselves with great diligence to the speaking of French in order to be thought more highly of..."
 (Mossé 1968, my translation)

Higden's statement, in addition to the fact that royal proclamations, laws, and parliamentary decrees were written in Norman French, indicates that there was a period in which a type of "diglossia"⁸ existed in England. That is, the use of French was considered appropriate for more formal modes of communication, commercial correspondence, and literary usage, while the use of English was relegated to more "humble" modes of communication, for example, the delivery of sermons to the common folk.⁹

There is abundant evidence that the usage of English became more extensive in the Thirteenth Century. It is telling that literature intended for the upper classes began to appear more regularly in English dialects, and that the authors often included either French or English translations of words and phrases used in their texts. For example, in Ancrene Riwe (ca. 1225), a treatise on religious life intended for aristocratic women entering convents, the following English clarifications of French expressions are included (Jespersen 1968:89):

"...cherité, þet is luve (charity, that is, love)..."
 "...ignoraunce, þet is unwisdom & unwitenesse
 (ignorance, that is, un-wisdom and un-wittingness)..."

La3amon's Brut (ca. 1200), a courtly romance written in English, contains the following French translations of English expressions (Jespersen 1968:89-90).

"...twelfe iferenen, þe Freinsse heo cleopeden
 duze pers (twelve companions, which [in] French
 they call "duze pers")..."
 "þat craft: to lokie in þa lufte, þe craft
 is ihote astronomie in oþer kunnes speche
 (that craft: to gaze into the sky; which craft
 is called "astronomy" in another sort of speech)..."

Baugh (1957) has amassed a large body of documentation which indicates that there was a language shift from French to English among the nobility during the late Thirteenth and early Fourteenth Centuries. The dating of French loans into English by Jespersen and Koszal (Jespersen 1968:87) shows clearly that the greatest influx of French lexical items occurred between the years 1251 and 1400. In 1258, King Henry III issued a royal proclamation, The Provisions of Oxford, to the nobility in English as well as French. This was the first known use of English in royal communication following the Conquest. In the introduction to his important work Speculum Vitae (1325), William of Nassyngton declared:

"...In the English tongue I shall talk to you,
 If you will bear with me so long.
 No Latin will I speak or waste,
 But English, which men use most,

Which each man can understand,
 Who is born in England;
 For that language is most pronounced,
 As well among learned as uneducated...
 Both learned and uneducated, old and young,
 All understand the English tongue..."

It is even more strongly indicative of the language shift that in ca. 1285 Walter of Bibbesworth published a manual intended to teach French to the Children of English aristocrats, which enjoyed wide circulation. In 1332 Parliament issued a decree urging that "...all lords, barons, knights, and honest men of good towns should exercise care and diligence to teach their children the French language..." (Baugh 1957:166). French was by that time obviously not the native language of the younger generation of English aristocrats.

There were significant political and economic changes in the Twelfth and Thirteenth Centuries which conspired to bring about these changes in the sociolinguistic profile of England. Political disputes between the Central French and Anglo-Norman kings precipitated a gradual alienation of the English nobility from the French culture. A long series of wars ensued between France and England during which English landowners came increasingly to identify themselves as Englishmen rather than Frenchmen. This process of reidentification was hastened by the fact that after the loss of Normandy to the Central French Crown, both the English and French kings demanded, in 1244, that Norman nobles pledge allegiance to one crown or the other. Anglo-Norman landowners were thus forced to relinquish their property on one side of the English Channel.

It is also noteworthy that Central French eclipsed Norman French as the prestige dialect of French Culture. The fact that the French loanwords which came into English after ca. 1300 are from the Central French dialect (Baugh 1957) indicates that this shift in prestige was responded to even by the Anglo-Normans.

During the period when the nobility were abandoning French, there began a great transformation in the social and economic structure of England.¹⁰ One factor which led to this transformation was the wooltrade with the Continent. The great demand for British wool that developed in Thirteenth Century Europe had three direct effects on English society. First, it put money into the hands of the peasantry as well as the landowners. "bondsmen" (or peasant farmers) were free to raise sheep and sell their wool. Some of the land-owning nobility and churchmen (or "landlords") were willing to commute the work owed them by their peasants in exchange for rent payments, which freed some bondsmen for other pursuits. Secondly, the wooltrade created an opportunity for the peasants now freed from the soil to become wool merchants. The number of wool merchants grew large enough in the early Fourteenth Century so that they organized mercantile associations to protect and advance their common interests. The third effect which the wool trade had was that it led to the increased growth of the towns, which were important

as wool-collection points. London was a particularly important focal point for the wool trade, because it was there that the "Staple", or great wool-merchants' exchange was established.

A second great factor in this restructuring of English society was the catastrophic epidemic of bubonic plague that swept the Island in the years 1348 and 1349. At least one-third of the population of England died in the Black Plague, which resulted in a severe labor shortage on the great country estates and in the towns. This put the surviving peasantry in a bargaining position which they were quick to take advantage of. They were able to demand wages as well as more land for their own use in exchange for labor. They became so bold in their demands that the wage level soared as they were able to strike from time to time. Some peasants were able to accumulate large tracts of adjoining fields which had been left unattended by the death of their less fortunate neighbors. The rise in wages compelled some nobles and churchmen to lease some of their property to the more industrious farmers, who were thus able to employ laborers themselves. For the first time, many bondsmen were able to buy their freedom and own property, and a new class of landed commoners developed.

A third development with great consequences for English society was the rapid growth of the textile industry, which occurred as the European upper classes came to value English cloth and to demand a great supply of it. A variety of specialized crafts were needed to produce standardized-quality cloth in large quantities. Accordingly, there arose a new class of capitalist entrepreneurs to organize the manufacture of cloth in towns and villages. In the Fourteenth Century, the craftsmen in the towns organized themselves into guilds. The entrepreneurs became an important faction in the Parliament, for the nobility turned to them to finance the armies which they sent to loot France to replenish their own wealth.

Thus, by the end of the Fourteenth Century profound and rapid changes had occurred in English society. In less than three generations, a new wealthy middle class of landowning commoners, merchants and manufacturers had come into existence. Many of these men were able to afford an education for their sons, who became an important political force in the following generations. The evolution of the modern system of social stratification out of the old feudal order was well under way.

The elevation in status of the common Englishman engendered an elevation of the importance of his language as a medium of commerce and government. The Great Death had an additional consequence for the status of English. It brought about the replacement of French by English in the schools. Nearly three-fourths of the clergy, including those in the teaching orders, died in the epidemic and later recurrences of the plague, and they were largely replaced by English-speaking individuals. John of Trevisa (1385), in an addendum to his translation of Higden's Polychronicon (cited above), noted that the use of French in the schools was declining in favor of English, principally due to the efforts of two Oxford school-

masters, John Cornwall and Richard Penrich (Baugh 1957:179). This is reflected in the fact that explicit regulations were established at monastery schools, colleges and universities ordering the use of French or Latin among the students. At the same time, however, Parliament issued a decree (1362) requiring the use of English rather than French in the courts of law.

With the evaporation of the sharp linguistic distinction which had formerly existed between the feudal castes in England at exactly the time when a competitive middle class was emerging, it seems reasonable to assume that new linguistic distinctions arose to maintain the social distance between the upper class and the well-to-do middle class. Hodges (1964:131) observes that

"...In the more rigid stratification systems of the past... social-class position and style of life were more congruent and manifest. And when dislocations occurred, when occasional merchants enjoyed greater wealth or power than occasional aristocrats, a visibly different style of life was often the only manner in which impoverished noblemen could effectively confirm their superior status.

"Dislocations of this sort were especially rife when, during the 'commercial revolution', the balance of power in western Europe shifted from the disintegrating feudal estates to the burgeoning cities. Wealth and political influence flowed into the hands of the urban bourgeoisie..."

We can feel confident in assuming that any noticeable linguistic differences between the language of the upper class would have been ready targets for exploitation as markers of relative social status.

The Dialects of the Upper and Middle Classes of Medieval London.

In attempting to define more precisely the sociolinguistic variables which led to the early transitions toward Modern English pronunciation, it is important to consider the dialectal composition of London in the Fourteenth and Fifteenth Centuries. By the Thirteenth-Hundreds, London had developed into one of the important commercial centers of the Western World. It was by far the most important city in England, not only because it was the hub of economic activity, but also because the Parliament was established there. Anglicists have long agreed that London English eventually came to serve as the basis for the literary standard which emerged in the Late Fifteenth and Early Sixteenth Centuries (Wyld 1927; Baugh 1957; Jones 1972). As Wyld observed, however, the development of London English involved a mixing of regional dialects within which one must acknowledge the probable emergence of social variation of the type observed in urban centers today (Wyld 1927:140-143, 146-150).

Across a gap of six centuries, we cannot hope to establish with absolute certainty the phonetic details of the variations that existed in Early London English; nor can we pretend to be able to discern with absolute clarity the social stratification of phonetic

variables across classes as linguists have succeeded in doing with respect to modern urban communities (Labov 1966; Trudgill 1974; Wolfram 1970, *inter alia*). The phonetic details are greatly obscured by the inconsistency of Late Middle English spelling and the lack of phonetic descriptions of English from before 1569 (Hart's *An Orthographie*). Our picture of the sociolinguistic distribution of phonological variants is likewise vague, owing to the absence of prescriptive statements about pronunciation from English literature until 1531 (Elyot's *Gouernour*). It is, however, possible to draw inferences about the phonetic variation that was probable in London English at the time in question by considering the phonological details of the dialects which were juxtaposed in the capital by the socio-economic developments traced above. Philological studies and demographic investigation that have been undertaken enable us to infer which dialects were brought together, and how these dialects were associated with different classes.

Through an extensive examination of public records, Ekwall (1956) has established that the middle class of Fourteenth and Fifteenth-Century London was composed largely of immigrants from the East Midlands, which was the great wool-producing area. Furthermore, his study shows that East Midland people occupied many influential positions as clerks, lawyers, pleaders, judges, public officials and parish priests (p. LXIII). It is therefore highly probable that the variety of speech most characteristically associated with the successful middle class had East Midland features.

There were, of course, other dialectal influences on the language of London. Wyld (1927:140-142) has found that the indigenous city dialect was probably Southeastern in type in the Thirteenth Century. But the literary language of the Fourteenth and Fifteenth Centuries shows great influence from the Central Southern dialect area. It seems reasonable to assume that this was due to the fact that the Royal Court in Middlesex and Oxford University, which was the primary center of learning, were both located in that dialect region. The language of royal proclamations and the language of London city documents alike exhibit Central Southern inflections and orthographic conventions (Chambers and Daunt 1931). The Fifteenth-Century *London Chronicle* also shows this type of dialect, even in the portion known to have been written by an East Midlander, Mayor William Gregory (Kjerrström 1946:17-18). The language in the works of Geoffrey Chaucer (1340-1400), the Court poet, is of this Southern variety as well, although his rhymes indicate a certain amount of phonological interference from the indigenous speech of the City (Wyld 1927:94, 109; Baugh 1957: 233). It is therefore highly likely that the variety of speech most strongly associated with the upper classes had the phonological characteristics of the Central Southern dialect.

Phonetic Variation in the Speech of Late Medieval London.

A comparison of the phonological systems of the juxtaposed

East Midland and Central Southern dialects reveals that they differed principally with respect to those lexical items that contained front vowels and upgliding diphthongs with frontal nuclei. Apparently, the vocalic systems of the two dialects were virtually identical with regard to the number and quality of their phonetic contrasts. Figure 3 represents the probable long vocalic system shared at the end of the Thirteenth Century, based upon the available knowledge about phonological developments in ME. (Wyld 1927, 1937; Mossé 1968; Prins 1974).

i:	ɪj	iu	ui	u:
e:	ɛj			o:
ɛ:	æj	ɛy	ɔj	ɔ:
		a:	ay	ɔy

Figure 3.

Figure 3 involves the assumption that the palatalized velar fricative [ɣ] (< Pre-English *[g]) and the voiceless palatal fricative [ç] (< Pre-English *[x]) had already by Late ME times been vocalized and merged as a palatal approximant [j]. The orthographic evidence for this vocalization dates from the Early Fourteenth Century (Prins 1974:76). Figure 3 also entails the assumption that the quantity distinction had been neutralized in the nuclei of upgliding diphthongs. There seems to be no orthographic evidence that contradicts this assumption.

It should be mentioned in passing that the exact qualities of the diphthongal nuclei of Late ME are unsure. The diphthong represented here as [æj] developed from OE [æg] and [ɛg] early in the ME Period, and was alternatively spelled ay and ey (Prins 1974:91). Traditionally, it has been assumed that these spellings indicate that the nucleus of [æj] was retracted in London English, and that there was a merger of these diphthongs as [aj]. But it is equally likely that this merged diphthong was [æj], since it is very likely that the letter a represented [æ] in isolation. What is absolutely clear from the spelling evidence is that this diphthong was different from the diphthong represented here as [ɛj], because the spellings ay, ey were used for the former, while the spellings egh, e3, eigh, ei3, eygh, ey3, eh and ey were used to represent the latter. This diphthong, in turn, was kept distinct from the diphthong represented here as [ij], which was consistently spelled in ME as ygh, igh, y3, i3, ye.¹¹ In light of the fact that the sequences [ix], [ixt], [ext] occurred in closed syllables in ME and the sequence [æx] occurred in antepenultimate syllables or closed syllables prior to vocalization, it is likely that the nuclei of the upgliding diphthongs were lax (Cf. Table 1).

Although their phonetic vowel systems were congruent, the dialects in question diverged with respect to the distributions of these segments in their lexicons. These divergences were the result of historical differences in the instantiations of the processes of breaking, palatal umlaut, lowering and raising. The correspondences of Table 2 illustrate these distributional differences.

Table 2. Correspondences between the front vowels and front diphthongs of Central-Southern and East-Midland Middle English, ca. 1300.

<u>Pre-English</u>	<u>Central-Southern</u>	<u>East-Midland</u>	<u>Modern Reflex</u>
*rixt	B,L,S,G rɛjt	B,L,S,G rɛjt	'right'
*gesixiθu	B,U,S,G rijt	B,S,G sijt	'sight'
*gewixiθu	B,U,S,G wijt	B,S,G wijt	'weight'
*sixiθa	B,U,S,G sijθ	B,S,G sijθ	'sees'
*fixtiθa	B,U,S,G fijt	B,S,G fijt	'fights' (V)
*li:xt	B,L,S,G lɛjt	B,S,G lijt	'light' (Adj)
*li:xtira	B,U,S,G lijter	B,S,G lijter	'lighter'
*li:xtjan	B,U,S,G lijten	B,S,G lijten	'to lighten'
*li:oxt	L,S,G lɛjt	S,G lijt	'light' (N)
*li:oxtjan	U,S,G lijten	S,G lijten	'to lighten'
*knext	B,R,S,G knijt	B,R,S,G knijt	'knight'
*knextas	B,S,G knɛjtes	B,S,G knɛjtes	'knights'
*fextan	B,S,G fɛjten	B,S,G fɛjten	'to fight'
*sexan	B,S,C se:n	B,S,C se:n	'to see'
*hæ:x	B,S,G hɛj	R,B,S,G hɛj	'high'
*næ:x	B,S,G nɛj	R,B,S,G nɛj	'nigh'
*hæ:xira	B,U,S,G hijer	R,B,S,G hɛjer	'higher'
*hæ:xista	B,U,S,G hijest	R,B,S,G hɛjest	'highest'
*mæxt	B,S,G mɛjt	B,S,G mɛjt	'might' (N)
*næxt	B,S,G nɛjt	B,S,G nɛjt	'night'
*æxta	B,S,G ɛjt	B,S,G ɛjt	'eight'
*mæxtig	B,U,S,G mijt	B,U,G mɛjti	'mighty'
*brigdel	G brijdel	G brijdel	'bridle'
*nigon	G nijn	G nijn	'nine'
*bugjan	U,Ur,G bijen	U,Ur,G bijen	'to buy'
*de:gen	G dɛjen	G dɛjen	'to die'
*le:gen	G lɛjen	G lɛjen	'to lie'
*fle:gen	G flɛjen	G flɛjen	'to fly'
*ɛ:agen	S,G ɛjen	S,G ɛjen	'eyes'
*hɛ:arjan	U,S hi:ren	U,S he:ren	'to hear'
*hɛ:r	he:r	he:r	'here'
*geldan	B,Lg ji:lden	Lg je:lden	'to yield'
*skeld	B,Lg ji:ld	Lg je:ld	'shield'
*swi:n	swi:n	swi:n	'swine'
*blind	Lg bli:nd	Lg bli:nd	'blind'
*mu:siz	U,Ur mi:s	U,Ur mi:s	'mice'
*dæg	G(R) dæj(dɛj)	G(R) dæj(dɛj)	'day'
*mæg	G(R) mæj(mɛj)	G(R) mæj(mɛj)	'may'

Table 2. (continued)

*weg	G(L)	wɛj(wæj)	G(L)	wɛj(wæj)	'way'
*regn	G(L)	rɛjn(ræjn)	G(L)	rɛjn(ræjn)	'rain'
*græ:g	G(R)	græj(grɛj)	G(R)	græj(grɛj)	'grey'
*dæ:d	(R)	dæ:d(dɛ:d)	(R)	dæ:d(dɛ:d)	'dead'
*stelan	Lg	stɛ:len	Lg	stɛ:len	'to steal'

In this table, the upper-case initials refer to the historic vowel changes which occurred to produce the Middle English forms represented:

- B = Breaking, i.e., diphthongization of a front vowel before a velar fricative (and of a mid front vowel following a palatal in the Southern dialect).
 L = Lowering of a diphthong nucleus.
 R = Raising of a diphthong nucleus.
 S = Smoothing, i.e., monophthongization of a diphthong.
 U = Palatal Umlaut, i.e., fronting (and/or raising) of a back vowel before a syllable containing [i] or [j].
 Ur = Unrounding.
 Lg = Lengthening (Cf. Table 1).
 G = Glide Formation, i.e., vocalization of [ɣ] from [g].

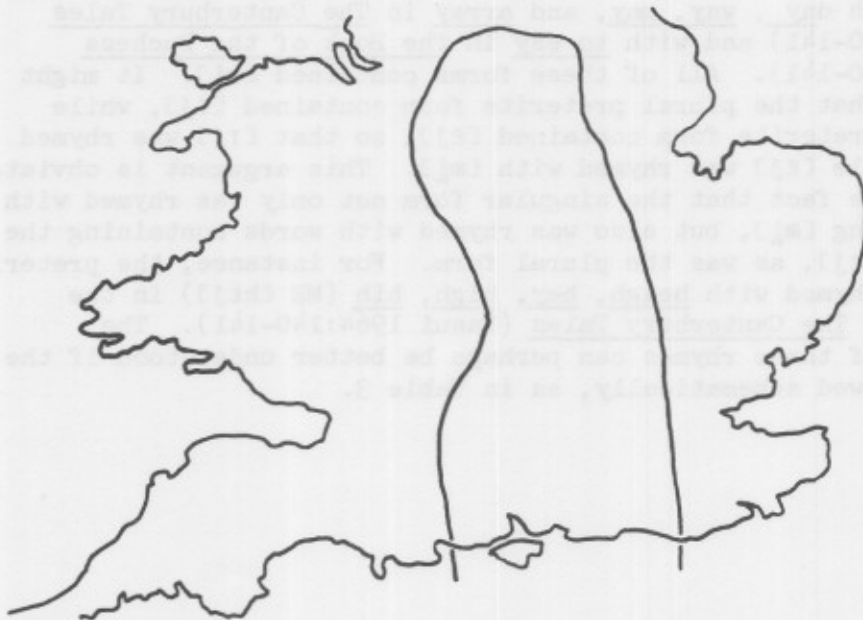
Central Southern (CS) has [ɛj] where East Midland (EM) has [ij] in lexical items containing the reflex of merged Pre-English (PE) *[i:x] and *[i:ox] when not in the position for palatal umlaut. This difference arose because OE [i:ɔ]--both derived from PE *[i:] by breaking before *[x] and directly inherited--was lowered in CSOE ("West Saxon") but not lowered in EMOE ("Anglian"). The resulting CS [e:ɔ] and EM [i:ɔ] were subsequently smoothed to [e:] and [i:], respectively. At the same time, CS and EM agree in having [ij] in those lexical items reflecting PE *[i:x] and *[i:ox] in position for palatal umlaut, i.e., preceding *[i] or *[j] in the next syllable.

CS also has [ij] in contrast to EM [ɛj] in lexical items reflecting PE *[æ:x] and *[æx] positioned for umlaut. In CSOE, breaking, umlaut and smoothing produced *[ɛ:ax]/[ɛax] > [i:ɛx]/[iɛx] > [i:x]/[ix], while in EM OE, umlaut raised PE *[æ:x] and *[æx] to *[e:x] and *[ɛx], which were broken, then smoothed back to EM [e:x]/[ɛx]. At the same time, however, CS and EM agree in having [ɛj] as the reflex of PE *[æ:x] and *[æx] in those lexical items which did not undergo palatal umlaut. Also, [ɛj] is shared as the reflex of PE *[e:g] and *[ɛ:ag].

Finally, CS had [i:] contrasting with EM [e:] in some common lexical items. In CSOE, *[ɛ:a] was unlauded to [i:ɛ] and later smoothed to [i:], whereas unlauded *[ɛ:a] gave [e:] in EMOE, which was not further altered. The vowel [i:] also resulted in CS from the breaking of PE *[e:] to [i:ɛ] following palatalized obstruents, with subsequent smoothing. In EMOE, this development did not occur. Once again, there were also lexical items which agreed in having [e:] as the reflex of PE *[e:] and [i:] as the reflex of PE *[i:], unlauded *[u:] and lengthened *[i] (Cf. Table 1).

Spellings from the first half of the Fourteenth Century indicate that there was a lowering of the diphthong [ɪj] to something phonetically close to [ɛj] in the area within the thumb-shaped isogloss marked on Map 1 (Serjeantson 1927; Wyld 1927; Oakden 1930; Prins 1974). Moreover, Serjeantson's survey of place-name spellings (1927) shows that the spellings indicative of this lowering were statistically most prevalent and earliest attested within the Central Southern counties. This suggests strongly that the lowering of [ɪj] began first in that area. The diphthong [ɛj], which had formerly been written with alternative spellings in which the nucleus was represented by e, came to be written in this area with spellings in which the nucleus was represented by i or the equivalent y. Thus the spellings igh, ygh, iȝ, yȝ, and ye came to be used to represent the historical developments of both ME [ɪj] and ME [ɛj] in the Central Southern area. This spelling convention was spread in the second half of the Fourteenth Century to London, where it appears in the manuscripts of Chaucer's works (Prins 1974). Later, it appeared in the London trade-guild documents (Chambers and Daunt 1931), official chronicle (Kjerrström 1946) and mercantile correspondence of the middle-class Cely Family (Malden 1900). In addition, it appeared outside London in the correspondence written by the educated members of the middle-class landowning Paston Family of Norfolk, in the East Midlands (Davis 1971).

Map 1. Area of Early 14th-Century i/y-spellings in lexical items with etymological [ɛj] from OE ēȝe/ēag, (after Oakden 1930).



The traditional interpretation of these spellings, proposed by Serjeantson (1927), Wyld (1927), Oakden (1930) and Prins (1974), is that they straightforwardly represent a raising of the nucleus of the diphthong [ɛj] in the central counties. Wyld and Prins go on to hypothesize that the raised pronunciation was spread into London English. Thus the reflexes of ME [ɛj] and ME [i:] were supposedly merged in the prestige dialect, from which the merged pronunciation [i] was eventually disseminated into the other dialects. In support of this hypothesis, Wyld (1927), following Wild (1915), asserts that Chaucer, reflecting the prestige dialect, consistently rhymed words containing the reflex of ME [ɛj] with words containing the reflex of ME [i:], although his scribes uniformly used the "traditional" spelling conventions which represented the nucleus of the diphthong as e. Prins (1974) adheres to Wyld's contentions about Chaucer's rhymes, although he recognizes the fact that Chaucer's scribes actually used alternative spellings to represent ME [ɛj], in which the nucleus could be written with i/y, as well as with e.

However, the detailed analysis of Chaucer's rhymes contained in Masui 1964 reveals that Chaucer actually rhymed words containing the reflex of ME [ɛj] with words containing [æj] as well as with words containing [i:]. For example, the nominative plural eyen ~ yen 'eyes' (< PE *[ɛ:agen])--which was rhymed with the infinitive deyen ~ dyen 'to die' (< PE *[de:gen])--was rhymed with the French loans espyen ~ aspyen ~ spyen 'to spy' (< OF espier) and cryen 'to cry' (< OF crier), both of which contained ME [i:], in Troilus and Criseyde (Masui 1964:141). The same form, eyen, was also rhymed with the Southern preterite plural form seyen '(they) saw', pronounced [sɛjen], in The Book of the Duchess (Masui 1964:141). The singular preterite form, variously spelled seigh, sigh, sy, say, was rhymed with day, way, may, and array in The Canterbury Tales (Masui 1964:140-141) and with to say in The Book of the Duchess (Masui 1964:140-141). All of these forms contained [æj]. It might be suggested that the plural preterite form contained [ij], while the singular preterite form contained [ɛj], so that [ij] was rhymed with [i:], while [ɛj] was rhymed with [æj]. This argument is obviated, however, by the fact that the singular form not only was rhymed with words containing [æj], but also was rhymed with words containing the reflex of ME [ɛj], as was the plural form. For instance, the preterite singular was rhymed with heigh, hey, high, hih (ME [hɛj]) in the manuscripts of The Canterbury Tales (Masui 1964:140-141). The implications of these rhymes can perhaps be better understood if the rhymes are viewed schematically, as in Table 3.

Table 3. Rhymes of words from the works of Chaucer. Lines indicate attested rhymes (based on Masui 1964).

<u>Reflecting ME [i:]</u>	<u>Reflecting ME [ɛj]</u>	<u>Reflecting ME [æj]</u>
	deyen ~ dyen	
espyen ~ aspyen ~ spyen	eyen ~ yen	
cryen	seyen;	seigh ~ sigh ~ sy ~ say—day, way, may, say, array
	heigh ~ high ~ hih ~ hey	

It is clear that the situation in Chaucer's speech cannot be as easily explained as has been traditionally assumed. The fact that Chaucer rhymed words containing the reflex of ME [ɛj] (henceforth referred to as "(ej)-class words") with words containing the reflex of ME [æj] (henceforth designated as "(æj)-class words") strongly suggests that Chaucer used, or was familiar with a pronunciation for (ej)-class lexical items in which there was a diphthong with a non-high nucleus phonetically similar to [æj], most probably a retention of ME [ɛj]. The rhymes of (ej)-class words with words containing the reflex of ME [i:] ("(i:)-class words") are open to three explanations. One possibility is that Chaucer was familiar with, or possibly used, a variant pronunciation for (ej)-class lexical items in which the nucleus was raised to [i] and thus rhymable with the nucleus of (i:)-class words. The second possibility is that Chaucer was familiar with, or used, a variant pronunciation for (i:)-class words in which the syllable nucleus had been diphthongized and lowered to something phonetically close to [ɛj]. The third possibility is that the vowel of (i:)-class words had been categorically changed to a diphthong phonetically close to [ɛj] in Chaucer's dialect. The alternative analyses envisioned are compared graphically in Figure 4.

<u>Alternative One:</u>	<u>Alternative Two:</u>	<u>Alternative Three:</u>
(i:)--[i:] (or [i̠])	[i:] ~ [ɛj]	[ɛj]
(ej)--[ɛj] ~ [ɪj]	[ɛj]	[ɛj]

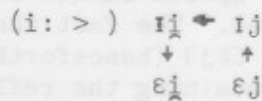
Figure 4. Alternative analyses of the rhymes in the works of Chaucer.

Alternative One provides partial support for the traditional hypothesis of Wyld and Prins, if we assume the Chaucer's rhymes represent an intervening period when the pronunciation of (ej) was variable before it was categorically changed to [ɪj] and subsequently merged with [i̠] ([i:]). However, this interpretation forces us to believe that the diphthong [ɛj] was first raised to [i̠], then later lowered (along with the reflex of ME [i:]) back to [ɛj], as the Vowel

Shift moved it toward [aj̄]. This hypothesis seems to be motivated principally by the fact that it allows all Modern English cases of [aj̄] to be uniformly traced back to earlier [īj̄].

Alternative Two is equally feasible, if not more feasible. It can be explained, too, as an intermediate stage of development extant in the speech of Chaucer's time while the reflexes of ME [i:] and [ij] were being lowered to merge with [ɛj]. In the end, the situation depicted in Alternative Three would have resulted. This hypothesis does not entail believing that ME [ɛj] was raised to [īj̄] and subsequently relowered to [ɛj̄]. Figure 5 depicts these alternatives schematically.

Traditional Hypothesis



Alternative Hypothesis

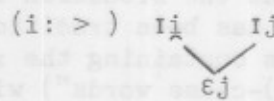


Figure 5.

The second hypothesis is equally supported by the i/y-spellings for (ej)-class lexical items if we regard these as "reverse spellings". That is, we can explain the spellings ygh, igh, etc. as having been extended to lexical items containing the diphthong [ɛj] because these spellings had been retained as conventional representations for the reflex of ME [ij], even after it had been lowered to [ɛj] or something phonetically close to [ɛj]. Thus the spellings with y/i were available as orthographic representations of [ɛj] regardless of its etymological origin.¹²

It appears from the evidence that the spellings ygh, igh, etc. were not used to write the reflex of ME [i:] in the Central Southern area in the first half of the Fourteenth Century, but confined to use for representing the reflex of ME [ij]. This indicates that in that dialect the reflexes of ME [i:] and [ij] were still phonetically distinct. That is, the lowering to [ɛj] affected the reflex of [ɛj] but not the reflex of [i:]. We cannot dismiss from consideration the possibility that the lowering of ME [ij] was influenced by phonological interference from the vocalic system of Anglo-Norman French. As has been observed above, there were no doubt many influential individuals at Oxford and the Royal Court in the Early Thirteen-Hundreds whose primary language was Anglo-Norman, but who also spoke English. Price (1971) reports the front-vowel system diagrammed in Figure 6 for Anglo-Norman. It is obvious from Figure 6 that the front-vowel systems were qualitatively quite similar, except for the fact that Anglo-Norman had no upgliding diphthong with a high front nucleus. The closest phonological element available for substitution by speakers whose primary language was Anglo-Norman was the diphthong [ɛj].¹³

Anglo-Norman	CSME
i	i: i ij
e εi	e: e εj
ε æi (εi)	ε: ε æi (εi)

Figure 6. Anglo-Norman and CSME phonetic front-vowel systems.

Regardless of the validity of this speculation, the fact remains that in the CS dialect the lowering came about in lexical items reflecting ME [ij]. The results of this change, applied to the examples of Table 2, are displayed in Table 4. This would, of course, have brought about a situation in London English where [ij] and [εj] were heard as variant pronunciations in (ij)-class lexical items. The dialects in contact there would have agreed in sharing the diphthong [εj] in some common lexical items, while they differed with respect to many other lexical items in which CS had [εj], whereas EM had [ij]. These differences in pronunciation were no doubt conspicuous, especially since they were correlated with differences in social status. Specifically, the pronunciation [εj] in (ij)-class words would have been identified with the upper classes and the educated, who were associated with the Central Southern dialect. The contrasting [ij] pronunciation, on the other hand, was likely identified with the middle-class, who, as we have seen, were probably associated with the East Midland variety of English.

Table 4. Correspondences between (ij) and (εj)-class words in the Central Southern and East Midland dialects, ca. 1350.

sijt; sijθ; fijt > sεjt; sεjθ; fεjt	sεjt; sεjθ; fεjt
wijt > wεjt	wijt
lejt	lijt
lijter; lijten	lijter; lijten
rεjt	rεjt
knijt > knεjt	knijt
knεjtes; fεjt	knεjtes; fεjt
hεj; nεj	hεj; nεj
hijer; hijest > hεjer; hεjest	hεjer; hεjest
miyti > mεyti	mεyti
dεjen; lεjen;	dεjen; lεjen;
flεjen	flεjen
εjen	εjen
biyen > bεyen	biyen

As we have also observed in the data of Table 2, these two varieties also differed in their phonetic realizations of certain lexical items containing the merged reflex of PE *[e:] and *[ε:a], while they agreed in the case of other lexical items. Thus,

correspondences of the types shown in Table 4 existed for what I will label (e:)-class words.

Table 5. Correspondences between (e:)-class lexical items in the Central Southern and East Midland dialects, ca. 1350.

<u>Central Southern</u>	<u>East Midland</u>
hi:ren	he:ren
ji:lɔn	je:lɔn
ji:lɔ	je:lɔ
te:θ	te:θ
he:r	he:r

Hyperextension and the beginnings of the Great Vowel Shift.

These considerations lead to the conclusion that the phonetic variants of (ij)-class and (e:)-class words would have been likely candidates for exploitation as phonological markers of social status within the sociolinguistic context that has been reconstructed for Fourteenth-Century London. The hypothesized social stratification of the variables (ij) and (e:) is summarized in Figure 7.

<u>Upper Class</u>	<u>Middle Class</u>
(ij): [ɛj]	[ij]
(e:): [i:]	[e:]

Figure 7. Social stratification of (ij) and (e:) in the English of 14th-Century London.

If the hypothesized sociolinguistic variation had truly existed, one would expect that the upwardly-mobile class of speakers would have tended to adopt the variants associated with higher social status, at least in some speech styles. As mentioned above, the *i/y*-spellings for (ej)-class words which are indicative of the lowering of [ij] are plainly evidenced in London writings associated with the middle class by the beginning of the Fifteenth Century. Examples of such spellings from London documents from 1384-1428 (Chambers and Daunt 1931) and the London Chronicles from before 1467 (Kjerrström 1946) are given in Table 6.

Table 6. Spellings of (ej)-class words from middle-class writings of 15th-Century London.

For [nɛj]	'nigh'	: <u>nigh</u>	beside: <u>negh</u> , <u>neigh-</u> , <u>neygh-</u>
[hɛj]	'high'	: <u>high</u>	<u>hei3e</u> , <u>hey(e)</u>
[nɛjt]	'night'	: <u>night</u> , <u>ny3t</u>	
[rɛjt]	'right'	: <u>right</u> , <u>ryht</u> , <u>ry3t</u>	

In addition, the (ij)-class lexical item [wɪjt] is found spelled weyght(e) as well as wyght, whyghte, whyte in the London documents (Chambers and Daunt 1931). There is also orthographic evidence of the same sort that indicates that the [ɛj]-variant spread outside the London area in the Fifteenth Century. Table 7 contains examples taken from the autograph letters of the Paston Family of Norfolk (Davis 1971).

Table 7. Spellings of (ej)-class words found in middle-class East Midland writings of the 15th Century.

For [mɛjtɪ]:	<u>myghty</u>	(Wm. Paston I)
	<u>mighti</u>	(John Paston I)
	<u>myty</u>	(Wm. Paston II)
[knɛjtəs]:	<u>knythys</u>	(Wm. Paston II)
[rɛjt]:	<u>ryght</u> , <u>ryth</u>	(John Paston I)
	<u>ryth</u>	(Wm. Paston II)

In essence, I am proposing that middle-class speakers who were upwardly mobile added to their grammars "adaptive rules" of the type described in Andersen 1973:773 ff. Such rules have the effect of modifying the output of a speaker's native phonological component in certain social contexts. For the case in point, the adaptive rules could be formalized as in Figure 8.

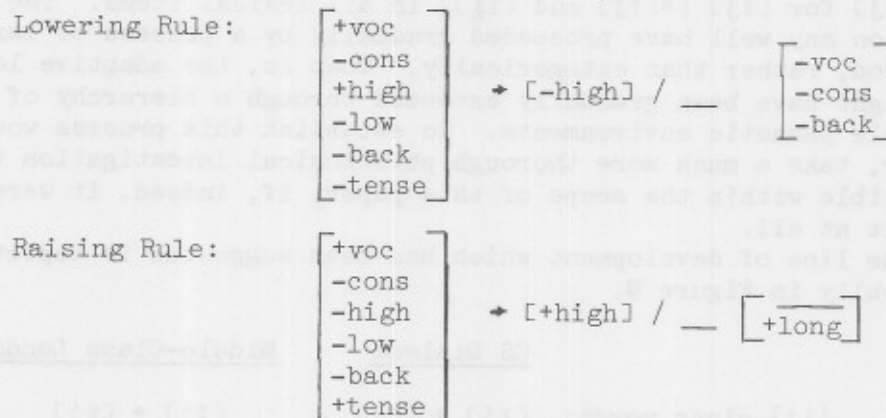


Figure 8. Adaptive rules of Lowering and Raising proposed for middle-class speakers of 15th-Century English.

It has been mentioned already that recent research concerning contemporary sound changes by Labov and others has shown that there is a tendency among upwardly-mobile speakers of lower social status to "hypercorrect", i.e., hyperextend the phonetic variants which they identify with higher social standing. In the situation described, one would expect that the addition of adaptive rules to the grammar would result in precisely this kind of hyperextension. This follows

from the fact that the speakers who were attempting to affect the phonological characteristics of another dialect would surely experience some uncertainty about the lexical environments in which the "target" variants would actually be used by the native speakers of the emulated speech variety. Again, there is orthographic evidence from the middle-class writings of the Fifteenth Century which indicates that this type of hyperextension did in fact occur.¹⁴

The i/y-spellings which were used to represent the diphthong [ɛj] are found in the writings of the Paston Family in lexical items that contained etymological [i:]. For example, the word 'write' (< OE [wri:tan]) is spelled wryth, wryth by John Paston I and William Paston II (who consistently metathesized h--their version of gh--and t, e.g., in ryth 'right'). This can be taken to indicate that the reflex of ME [i:], which had very possibly been diphthongized to [ij] by some speakers, was phonetically identified with the diphthong [ɛj] by some and thus "incorrectly" lowered to [ɛj] just as [ij] was. From the point of view of those speakers who perceived and pronounced the reflexes of ME [ij] and [i:] as phonetically equivalent, it would have been logical to hyperextend the variant [ɛj], as a marker of higher status, to all lexical items which they perceived as containing the socially "inferior" phonetic variant [ij]. Such speakers could not, of course, have been aware of the differing historical origins of the phonetic segments which they perceived as merged. This hyperextension would have had the effect of substituting [ɛj] for [ij] (= [ij] and [ij̄]) in all lexical items. The hyperextension may well have proceeded gradually by a process of lexical diffusion, rather than categorically. That is, the adaptive lowering rule might have been gradually extended through a hierarchy of favorable phonetic environments. To establish this process would, however, take a much more thorough philological investigation than is possible within the scope of this paper, if, indeed, it were possible at all.

The line of development which has been suggested is depicted graphically in Figure 9.

	<u>CS Dialect</u>	<u>Middle-Class London Dialect</u>
(ij)-class words:	[ij] → [ɛj]	[ij] → [ɛj]
(i:)-class words:	[i:]	[ij̄] → [ɛj]

Figure 9. Hypercorrective extension of the variant [ɛj] in Early Modern English.

If this type of development had actually occurred, one would expect to find examples in Fifteenth-Century writings where the spelling y/i, traditionally used to represent [i:], was employed to represent the diphthong [ɛj] or diphthongs phonetically similar to [ɛj] as a reverse spelling. And, in fact, such spellings are attested in the London documents already cited (Chambers and Daunt 1931).

Table 8. y/i-spellings for (ej) and (æj)-class lexical items in 15th-Century London English.

For	[neɟ]	'nigh'	:	<u>ny-</u>
	[lɛjɛr]	'liar'	:	<u>lyere</u>
	[sæjnt]	'saint'	:	<u>synt</u>

Also there appear phonetic spellings in which the reflex of ME [i:] is written ey, a spelling adopted in the Later ME Period for (ej)-class words, when gh was replaced by y in many manuscripts. These spellings are exemplified in Table 9.

Table 9. ey-spellings for (i:)-class lexical items found in 15th-Century English documents.

For ME	[pi:nɛn]	'to pine'	:	<u>peynen</u>	(Kjerrström 1946)
ME	[twi:s]	'twice'	:	<u>tweys</u>	(Chambers and Daunt 1931)

The same reasoning would lead us to expect that the higher-status variant [i:] for (e:)-class words would have been hyperextended to all (e:)-class lexical items eventually, regardless of their etymologies. In this instance, too, upwardly-mobile speakers who were adapting their phonological output would have been unaware of the historical differences between those lexical items in which the higher-status variety had [i:] in contrast to their [e:], and the lexical items in which both varieties had [e:]. To them it would merely have been conspicuous that higher-status speakers often said [i:] where they said [e:].

Fifteenth-Century writings again contain orthographic evidence of the hyperextension. In Table 10 examples are given in which the spelling y, traditionally used for (i:)-class words, is employed to represent the nucleus of (e:)-class words. These examples are drawn from the section of the London Chronicles written by Major William Gregory (Kjerrström 1946) and from the correspondence written for the Paston Family by an estate employee named John Wykes (Davis 1971).

Table 10. y-spellings in (e:)-class words taken from middle-class writings of the 15th Century.

For	[agre:d]	'agreed'	:	<u>agryed</u>	
	[de:meθ]	'(he) deems'	:	<u>dymeth</u>	
	[he:r]	'here'	:	<u>hyre</u>	(John Wykes)
	[ke:p]	'keep'	:	<u>kype</u>	
	[spe:d]	'speed'	:	<u>spyde</u>	
	[we:pɪŋg]	'weeping'	:	<u>wyping</u>	
	[de:r]	'deer'	:	<u>dyre</u>	
	[ste:pell]	'steeple'	:	<u>stypylle</u>	(William Gregory)
	[sle:vɛs]	'sleeves'	:	<u>slyvys</u>	

The addition of the adaptive rules of lowering and raising to the grammars of adult speakers of Early Modern English and their subsequent hyperextension would have led to a sociolinguistic context within which the natural processes of lowering and raising proposed by Miller (1972) and Stampe (1972) could be actuated in the speech of the younger generation, rather than suppressed as in the speech of earlier generations. The lowering process, shown in Figure 10, would have had the effect of substituting the pronunciation [ɛj] for the adult pronunciation [i:] in those dialects where the high upgliding diphthong existed in (ij)-class words. The raising process, also portrayed in Figure 10, would have had the effect of substituting [i:] for the adult pronunciation [e:]. In both instances, the substitutions produced by the operation of the proposed natural processes would have been precisely those phonetic variants which were more highly-valued in the social milieu of upwardly-mobile, more prestigious adult speakers.

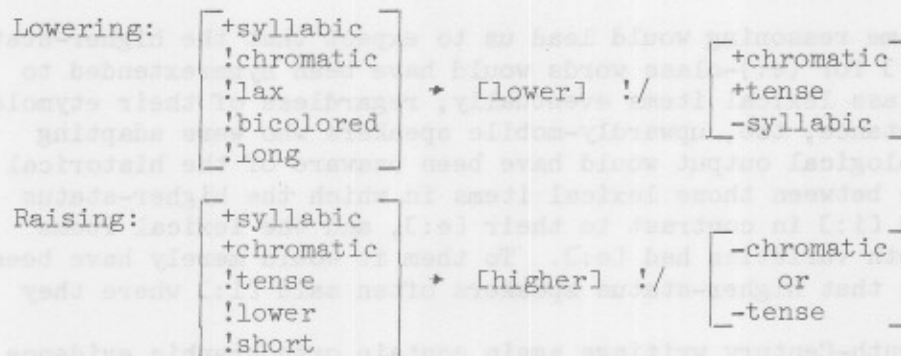


Figure 10. The natural processes of Lowering and Raising (based on Miller 1973).¹⁵

Concluding remarks.

This paper had dealt with only two of the series of changes involved in the Great Vowel Shift of English, by way of example. The same approach, however, could be applied to solve the transition and embedding problems for the other changes involved in the Vowel Shift. In fact, the later stages of the Vowel Shift could be even more easily analyzed because more overt sociolinguistic information and more transparent orthographic evidence is available with regard to them. The sociolinguistic variation involved in the raising of ME [ɛ:] to Modern English [i:] has been alluded to by Wyld (1927; 1933), Kökeritz (1953), Weinreich, et al. (1968), and Labov (1972). There is an obvious correlation between the changes of the front vowels and diphthongs and the changes of the back vowels and diphthongs which has not been dealt with here. However, the type of analysis which Labov (1966, 1972) proposed to explain the correlated raising of (eh) and (oh) in New York City English, viewing them as systematically connected developments, could plausibly be combined

with philological analysis to account for the concomitant developments of ME [i:] and [u:], and ME [e:] and [o:].

It has been shown that the contemporaneous changes in the social and linguistic profiles of England during the Fourteenth Century led to a situation where two regional dialects with sizeable numbers of speakers were juxtaposed and realigned as social dialects. Furthermore, it has been demonstrated that the different phonological histories of the convergent dialects provided the raw material for sociolinguistic variation at that particular point in time when older sociolinguistic barriers were disintegrating in England. Finally, it has been hypothesized that the type of cross-dialectal phonological restructuring that has been observed in contemporary settings would have led to the initial phase of the Great Vowel Shift of English as a matter of course, given the phonetic variation and socio-dialectal alignment pattern that have been reconstructed for Early Fifteenth Century London English.

Footnotes

1. Based on Wright and Wright (1924).
2. Except when u preceded the labials m or p or followed w or y, giving the exceptions room < rūm, stoop < stoupe, droop < droupen, tomb < toumbe, cooper < couper(e), wound (noun) < wundian, you, your (in which ou was a ME spelling for [u:]). Cf. Prins (1974:130).
3. Except when e preceded the anterior stops or fricatives /d, t, θ, f, v/, in the following lexical exceptions: dread, breath, spread, wet, thread, sweat, shed, bread, dead, death, head, deaf, red, get, stead, heaven, tread, heavy, fret. Cf. Prins (1974:141).
4. The philological evidence bearing upon the individual changes has been succinctly assembled in Prins (1974) and the orthoepic evidence is critically reviewed in Wolfe (1972).
5. Stampe (1973) presents a more detailed picture of the theory of natural processes, and Miller (1973) relates this theory to the explanation of numerous context-free sound changes.
6. Thus, Wolfe's conclusions support the consensus of Horn (1908), Jespersen (1909), Luick (1900, 1914-1940), Ekwall (1914), Zachrisson (1927), Wyld (1937), Chomsky and Halle (1968), Stampe (1972) and Miller (1972). Labov, Yaeger, and Steiner's conclusions support the view shared by Ellis (1874), Sweet (1888) and Orton (1974). Proponents of hypothesis 1 are Dobson (1957), Stockwell (1952), McCawley (1969) and Bailey (1969). Proposal 2 is the analysis advanced by Kökeritz (1953) and Prins (1974).
7. Or trilingual, if Latin is assumed to have been widely spoken.
8. Or "triglossia", if it is assumed that certain types of communication were conducted exclusively in Latin.
9. Ferguson (1959) describes similar linguistic situations that exist in contemporary bilingual societies.

10. An excellent, detailed account of the development of English society is contained in Trevelyan 1942, from which the sketch presented here is drawn.

11. The characters i and y were used interchangeably for [i:] and [i] in Late OE and in ME, with a preference shown for y. The digraph gh and the letter ȝ were used interchangeably for post-vocalic [j]. In Late ME, y became an additional alternative spelling for [j].

12. This situation is perfectly analogous to the case of the "spurious" diphthongs of Classical Greek, discussed in Buck 1955 and Allen 1974.

13. Pope (1952) thinks it is possible that Anglo-Norman [ɛi] had already been monophthongized to [ɛ:] by the time in question, at least in the speech of some individuals; however, Price is confident that [ɛi] remained, at least in open syllables, i.e., the same position in which ME [ij] occurred after vocalization of [y].

14. Note that I am here assuming, like others who have dealt with cross-dialectal "borrowing" as a source of sound change, that adult speakers are capable of imitating the phonetic details of another dialect. As far as I know, this assumption has never been empirically investigated. Such an investigation is the subject of my forthcoming doctoral dissertation. Here the matter is not debated because there seems to be no reason to believe that the dialects under consideration differed significantly with respect to the phonetic segments in question.

15. In Miller's formalism, the symbol ! means "especially when", and the term "chromatic" means palatal or labial, i.e., front or rounded. "Bicolored" means both palatal and labial.

References

- Allen, W. 1974. *Vox Graeca*, 2nd ed. Cambridge: Cambridge University Press.
- Anderson, H. 1973. Abductive and deductive change. *Language* 49. 765-793.
- Bailey, C-J. 1969. The English great vowel shift past and present. *University of Hawaii Working Papers in Linguistics* 2.1-5.
- Baugh, A. 1957. *A History of the English Language*. New York: Appleton-Century-Crofts, Inc.
- Buck, C. 1955. *The Greek Dialects*, 3rd ed. Chicago: The University of Chicago Press.
- Cam, H. 1952. *England Before Elizabeth*. London: Hutchinson House.
- Chambers, R. and M. Daunt, eds. 1931. *A Book of London English 1384-1425*. Oxford: Clarendon Press.
- Chomsky, N. and M. Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row.
- Danielsson, B., ed. 1955. *John Hart's Works on English Orthography and Pronunciation*. Stockholm: Almqvist and Wiksell.

- Davis, N., ed. 1971. *Paston Letters and Papers of the Fifteenth Century*. Oxford: Clarendon Press.
- Dobson, E. 1957. *English Pronunciation 1500-1700*. Vol. 2: Phonology. Oxford: Oxford University Press.
- Dunn, C. and E. Byrnes, ed. 1973. *Middle English Literature*. New York: Harcourt Brace Jovanovich, Inc.
- Ekwall, E. 1914. *Historische neuenglische laut- und formenlehre*. Berlin: G. J. Göschen.
- Ekwall, E. 1956. *Studies on the Population of Medieval London*. Stockholm: Almqvist and Wiksell.
- Ellis, A. 1874. *Early English Pronunciation*. Vol. 4 (1968 reprint). New York: Greenwood Press.
- Ferguson, C. 1959. Diglossia. *Word* 15.325-340.
- Hodges, H. 1964. *Social Stratification*. Cambridge: Schenkman Publishing Co.
- Horn, W. 1908. *Historische neuenglische grammatik*. Strassburg: K. J. Trübner.
- Jespersen, O. 1909. *A Modern English Grammar on Historical Principles*. Part I. Heidelberg: Carl Winter's Universitätsbuchhandlung.
- Jespersen, O. 1948. *Growth and Structure of the English Language*. 9th ed. Oxford: Basil Blackwell.
- Jones, C. 1972. *An Introduction to Middle English*. New York: Holt, Rinehart and Winston, Inc.
- Kökeritz, H. 1953. *Shakespeare's Pronunciation*. New Haven: Yale University Press.
- Kjerrström, B. 1946. *Studies in the Language of the London Chronicles*. Uppsala: Almqvist and Wiksell.
- Labov, W. 1963. The social motivation of a sound change. *Word* 19. 273-309.
- Labov, W. 1966. *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.
- Labov, W., M. Yaeger and R. Steiner. 1972. *A Quantitative Study of Sound Change in Progress*. Philadelphia: The U.S. Regional Survey.
- Lass, R. 1974. Linguistic orthogenesis; Scots vowel quantity and the English length conspiracy. *Heslington: York Papers in Linguistics* 4.7-26.
- Loyn, R. 1967. *The Norman Conquest*. London: Hutchinson House.
- Luick, K. 1900. Der ursprung der neuenglischen ai- und au- diphthonge. *Englische Studien* 27.89-100.
- Luick, K. 1914-1940. *Historische grammatik der Englischen sprache*. Oxford: Basil Blackwell.
- Mac Farlane, K. 1973. *The Nobility of Later Medieval England*. Oxford: Clarendon Press.
- Malden, H., ed. 1900. *The Cely Papers: Selections from the Correspondence and memoranda of the Cely Family, Merchants of the Staple, A.D. 1475-1488*. London: Longmans, Green and Co.
- Martinet, A. 1955. *Économie des changements phonétiques*. Berne: A. Francke.

- Masui, M. 1964. *The Structure of Chaucer's Rime Words: An Exploration Into the Poetic Language of Chaucer*. Tokyo: Kenkyusha.
- Miller, P. 1972. Some context-free processes affecting vowels. Columbus: OSU WPL 11.
- Miller, P. 1973. Bleaching and Coloring. Chicago: *Papers from the Ninth Regional Meeting of the Chicago Linguistics Society*.
- Mossé, F. 1968. *A Handbook of Middle English*. J. A. Walker, trans. Baltimore: Johns Hopkins Press.
- Oakden, J. 1930. *Alliterative Poetry in Middle English*. Manchester: Manchester University Press.
- Orton, H. and E. Dieth. 1970. *Survey of English Dialects*. Vol. III: The East Midland counties and East Anglia. Leeds: D. J. Arnold.
- Pope, M. 1952. *From Vulgar Latin to Modern French with Especial Consideration of Anglo-Norman: Phonology and Morphology*. 2nd ed. Manchester: Manchester University Press.
- Price, G. 1971. *The French Language Present and Past*. London: Edward Arnold, Ltd.
- Prins, A. 1974. *A History of English Phonemes from Indo-European to Present-day English*. Leiden: Leiden University Press.
- Serjeantson, M. 1927. The development of O.E. *ēah*, *ēag* in M.E. *Journal of English and Germanic Philology* 26.198-230; 350-400.
- Stampe, D. 1972. On the natural history of diphthongs. Chicago: *Papers from the Eighth Regional Meeting of the Chicago Linguistics Society*.
- Stampe, D. 1973. *A Dissertation on Natural Phonology*. Unpublished University of Chicago Ph.D. dissertation.
- Stockwell, R. 1952. *Chaucerian Graphemics and Phonemics: A Study in Historical Methodology*. Unpublished University of Virginia Ph.D. dissertation.
- Sweet, H. 1888. *History of English Sounds*. Oxford: Clarendon press.
- Trevelyan, G. 1944. *English Social History: A Survey of Six Centuries, Chaucer to Queen Victoria*. London: Longmans, Green and Co.
- Trnka, B. 1959. A phonemic aspect of the great vowel-shift. In: *Mélanges de linguistique et philologie Ferdinand Mossé in memoriam*. Paris: Didier.
- Trudgill, P. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.
- Weinreich, U., W. Labov and M. Herzog. 1968. Empirical foundations for a theory of language change. In: W. Lehmann and Y. Malkiel, eds., *Directions for Historical Linguistics*. Austin: University of Texas Press.
- Wild, F. 1915. *Die sprachlichen eigentümlichkeiten der wichtigeren Chaucerhandschriften und die sprache Chaucers*. Vienna: W. Braumüller.
- Wolfe, P. 1972. *Linguistic Change and the Great Vowel Shift in English*. Berkeley: University of California Press.
- Wolfram, W. 1970. *A Sociolinguistic Description of Detroit Negro Speech*. Washington, D.C.: Center for Applied Linguistics.

- Wright, J. and E. Wright. 1924. *An Elementary Historical New English Grammar*. Oxford: Oxford University Press.
- Wyld, H. 1927. *A Short History of English*. London: John Murray.
- Wyld, H. 1933. *A History of Modern Colloquial English*. New York; E. P. Dutton and Co.
- Zachrisson, R. 1913. *Pronunciation of English Vowels 1400-1700*. Göteborg: Wald. Zachrissons Boktryckeri.

