ACQUISITION OF VARIABLE RULES: (-t,d) DELETION AND (ing) PRODUCTION IN PRESCHOOL CHILDREN

Julia Lee Roberts

A DISSERTATION

in

Linguistics

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

1994

____
William Labov
Supervisor of Dissertation

____
Donald A. Ringe
Graduate Group Chairperson

DEDICATION

This dissertation is dedicated
to the memory of
my grandfather

EDGAR LEE MARCRUM

1881 to 1962

whose belief in the value of education and the pursuit of knowledge
has continued to provide me with inspiration.

ACKNOWLEDGMENTS

It is hardly original to note that the process of writing a dissertation cannot be completed without the help and support of many people. This is particularly true when the dissertation has taken a considerable time to write, as has this one. I have many more people to thank than I could write on these pages, but a partial list is as follows:

The first thank you must go to my advisor, William Labov. His input into this dissertation began many years before my entrance into graduate school, when I was a practicing speech-language pathologist, and his work on Black English Vernacular taught me that there was far more to language than that which could be captured in a language evaluation measure. As his student and advisee, I have greatly benefitted from his generous sharing of insight. In addition, he has been unstinting in his support of my desire to combine my interests in sociolinguistics and child language. He has provided research assistant positions which have aided me financially as well as added to my knowledge of these subjects. Finally, and most recently, he has given me invaluable help in reaching what is undoubtedly the light at the end of the dissertation tunnel -- a "real job".

Gratitude must also be expressed to the rest of my rather large dissertation committee. Sherry Ash provided more hands-on help in this process than I would expect from an entire committee. She shared her prodigious knowledge in several areas from variable rule analysis and statistics to theoretical issues. In addition, she did not once express what would have been understandable frustration and annoyance at any of my increasingly frantic telephone calls and questions when I could not get the computer to do what I wanted it to. Finally, her friendship throughout my years at Penn has been irreplaceable, and I thank her for this with the hope that it continues for a long time to come.

Terry Pica stepped into my committee generously when I needed one more Penn faculty member to complete it. In addition, she provided valuable feedback on issues from another theoretical perspective. Finally, both Bambi Schieffelin and Gillian Sankoff have, through classes and as my committee members, provided me with an anthropological perspective which has informed my work thus far and and will continue influence it in future years. Both have also been excellent role models and given me personal and professional support, and I thank them for this.

Leslie Rescorla offered me a part time research coordinator position in the Psychology Department of Bryn Mawr College in the fall of 1992 and quite a bit more. Through her, I have gained insight into the field of psychology and its possible areas of intersection with linguistics. She has been generous with advice and support and extremely flexible in terms of time in allowing me to combine my obligations at Bryn Mawr with those at Penn and at home. She has also given me the opportunity to regain and add to my clinical skills in the area of speech pathology. Without this opportunity, it is doubtful that my job search would have been successful. Finally, she has allowed me access to her wonderful and painstakingly collected longitudinal data set, and I look forward to further collaboration with her in the future. My experience at Bryn Mawr has also been enriched by other faculty and staff members, and I want, in particular, to acknowledge Hollis Scarborough, Anne Fowler, Anne Marie Borneman and Nancy Rassiga.

My colleagues at Penn over the years have been many and valued. I cannot mention them all, but I would like to single out a few who have been especially helpful. Gayla Iwata-Reuyl has been a treasured friend and colleague, and, most generously, my son's first and favorite babysitter. Richard Cameron, Ann Taylor, Claudia Mazzie, Zhizhe Hwang, Maria Theresa Turrell, Diana Gonzolez, Vera Regan, Tom Veatch, Robin Sabino, Ruth Herold, Megan Moser, and many others have given me support, knowledge and friendship over the years.

Finally, and most personally, I would like to thank my family. My parents, Eugene and Eleanor Roberts, and my sisters, Jan Roberts-Breslin and Elizabeth Roberts, have all been patient, forbearing, and, most important of all, amusing and distracting, when I seemed incapable of amusement and driven to distraction. Several years ago, my husband, Bruce Brown, joined a club of which no one in his right mind wants to be a member -- the spouses of doctoral students club. He has fulfilled his membership duties well. I thank him for unending support during this time and look forward with him to his giving his position in the club to some other unsuspecting husband or wife. Lastly, I want to express my lasting gratitude to my children, Matthew and Annie Brown, both born during my doctoral studies.

They have provided me with inexpressible joy and have reminded me again and again of what is really important.

I know I am very fortunate to have had all of the help and support I have during the writing of this dissertation. Its inadequacies are, of course, my own, but all worthwhile in it is the result of the generosity of this large group of people. To all of them I am very grateful.

ABSTRACT

ACQUISITION OF VARIABLE RULES: (-t,d) DELETION AND (ing) PRODUCTION IN PRESCHOOL CHILDREN

JULIA LEE ROBERTS

WILLIAM LABOV

There have been many studies over the past few decades documenting the existence of variable rules in adult language. It is only recently, however, that the acquisition of these rules has been the focus of research, and that event has opened the door for questions about the interaction of the learning of categorical rules and that of variable rules. Specifically, questions have arisen as to whether these rules might not be construed as either a performance factor and/or a reflection of universal constraints on language.

The present study examines the acquisition of (-t,d) deletion and (ing) production in 3- and 4-year-old children in order to ascertain their degree of mastery of phonological, grammatical, and social constraints. Seventeen children were tape recorded during play interview sessions in their South Philadelphia day care center. Six to thirteen sessions per child over a three month period were required to obtain sufficient data for analysis. In addition, eight of their parents were interviewed in their homes for purposes of comparison.

Results of the study revealed that children as young as three had, for the most part, mastered the process of variation of (ing) and the phonological constraints on (-t,d) deletion, and they were well

into the process of acquiring the grammatical constraints on (-t,d) deletion. Their learning of a dialect specific phonological constraint demonstrated that their mastery of this variable rule was not a reflection of universal constraints. Further, their independent analysis of semi-weak verbs made it clear that they were not simply copying frequencies of their parents' forms but learning an abstract rule. The children's acquisition of the extralinguistic constraints on these rules lagged behind that of the linguistic factors. Of particular interest to the issue of gender differences in language was the girls' surprising tendency to delete (-t,d) more often than the boys, demonstrating that they had not yet learned linguistic conservatism in instances of stable variation and arguing against a biological basis for sex-based sociolinguistic differences.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.0 Background

It is the aim of this study to examine an area of language which has been often overlooked by both sociolinguists and psycholinguists -- the acquisition of variation by preschool children. An underlying assumption of this work will be that the knowledge that can be obtained from such a study is important, not just to one of these fields, but to both of them.

Weinreich, Labov, and Herzog (1968) argued that it is unrealistic to study language as a "homogeneous object" as it is conceptualized under traditional models of language description. They stated that "nativelike command of heterogeneous structures is not a matter of multidialectalism or 'mere' performance, but is part of unilingual linguistic competence" (p. 101). In the years which followed the publication of this work, many studies have documented the inherent variability in language, both in instances of language change and in cases of stable variation. Still, with very few exceptions, these studies have concentrated on the language systems of adult speakers. Rarely have they included speakers under the age of nine. The result is that the overwhelming majority of the studies on variation in language has concentrated on speakers well above the critical period for language acquisition.

The reasons for this exclusion of younger speakers from work on language variation are both theoretical and methodological and have changed over the last twenty years. Early work in this area tended to assume that before adolescence, there was little stylistic variation in children's speech, and that the peer group influence which took place in later childhood was of primary importance in the acquisition of variable rules (Labov, 1970; Wolfram and Fasold, 1974). Lakoff, in her 1973 work on women's language, asserted that both boys and girls began learning 'mommy's language' from their mothers and other early teachers and care givers, most of whom are female. As the children matured, their language became sex differentiated with boys most influenced by their peer groups to use 'male' language forms.

More recently, indications that children begin to acquire variable rules and some of their constraints very early have begun to emerge from research on variation (Kovac and Adamson, 1981; Labov, 1989a; Guy and Boyd, 1990), but methodological difficulties, specifically in the collection of sufficient data from children to allow for statistical analyses of variation in their language, have inhibited work in this area. It is more difficult as well as more time consuming to obtain sufficient data for individual variable rule analyses from children than from adults. The preferred solutions to this dilemma in the past have been to group children so that enough data is provided for meaningful statistical analysis or to plot the individual data on a graph but not to analyze it statistically. Kovac and Adamson grouped children by age, race, and social class resulting in groups of three to six children. Guy and Boyd grouped all of the children under 10 and some of those under 15 years old into neighborhood peer groups resulting in six groups of two or three speakers. Sankoff and Laberge (1973 [1980]) plotted individual data on the Tok Pisin future marker *bai*, including that of a 5-, 6-, and 8-year-old and three 11-year-olds ranging from 12 to 47 tokens each. Similarly, Labov listed the data for his single 4-year-old speaker but did not perform statistical analysis on it. While these procedures have resulted in important findings on the learning of (-t,d) deletion, grammaticalization of pidgins, (ing) production, and finite *be*, more data from each child is needed to enable researchers to match each individual speaker more conclusively to the speech patterns of the community as a whole.

The problem is compounded by the difficulty of catching the child early enough in her development of grammatical forms to be able to examine their earliest emergence while at the same time waiting until the child is verbal enough to produce adequate amounts of data for analysis. As will be discussed in detail in Chapter 6, one of the most important purposes of the present study will be to fill the gap created by this dilemma by developing strategies of collect large amounts of data from 3- and 4-year-olds for the purpose of individual and group quantitative analysis.

These first indications from the research presented above that acquisition of variation may begin during the preschool years when children are also acquiring the vast majority of categorical rules underline the importance of this work to the field of sociolinguistics. They suggest that, if these early findings are born out in larger, more comprehensive studies, the structured heterogeneity discussed by Weinreich et al. extends even to the youngest members of the speech community. Far from being a product of performance, as structural linguists might suggest, or even of multidialectalism, variation is embedded in the language acquisition process from its earliest stages.

Clearly, findings such as these would have relevance to the field of psycholinguistics as well. As has been the case in structural linguistics, this field has largely ignored variation in child language. Unlike in many generative linguistic studies of adult language which use judgements of the grammaticality of utterances, often made by the linguist him or herself, the data for child language studies have always been the actual speech of children. The forthcoming grammars, however, were based on, as Menyuk (1977) notes, "observations in the *structural consistencies* in these utterances" (emphasis added). Although certainly variation has been discussed, the types of variation which have been examined are usually restricted to three. The first is developmental variation which is brought about by changes in a child's language as he or she matures (Brown, 1973). Secondly, individual differences among children are often noted, but discussion of them is generally confined to those related to differences in cognition, as in studies of communication disorders, learning style, or caretaker input (Nelson, 1973, 1975; Bloom and Lahey, 1978). Finally, intrasubject differences as a result of environmental influences on the child are acknowledged. These are the ones often referred to as performance differences in generative linguistic literature.

It is clear from the above discussion, however, that the rule-governed variation which has been found time and time again to be a part of the language of adult speakers is also a part of the overall linguistic competence which a child must acquire in order to be a speaker of her language. Furthermore, as Andersen (1990) notes, "children must learn the dialect or set of dialects that will mark certain aspects of their social identity, including their region of origin, as well as their social class, ethnic group age, and gender" (p.32). In addition, they must learn the stylistic variation that will allow them to move from social group to social group, setting to setting, and topic of conversation to topic of conversation.

Andersen points out that it is necessary for researchers in child language acquisition to understand variation in order to have an accurate picture of normal patterns of development. She notes, for example, that to understand how children acquire negation, it is necessary to realize that doubly marked negatives may characterize an intermediary step for a speaker of one dialect (Standard English) but an endpoint for a speaker of another (African American Vernacular English). Andersen's point is a good one but does not go far enough. To acknowledge the acquisition only of non-variant language is to put forward an incomplete picture of language development. As previously noted, Weinreich et al. point out that to attempt to describe language as a homogeneous object is "unrealistic" and that it is necessary to learn to see language "as an object possessing orderly heterogeneity" (p. 100). The pursuit of a model of language acquisition which denies the presence, much less the importance, of the acquisition of this heterogeneity is similarly unrealistic. Rather, a complete acquisitional model demands the inclusion of all forms of language, those which are variable as well as those which are categorical in nature.

1.1 Statement of the problem

As previously stated, the overall aim of this study is to look closely at the acquisition of variation in children within the age range in which language is first acquired. To this end, the learning of the rule-governed variation found in (-t,d) deletion and (ing) production in 3- and 4-year-old children will be examined. This age range was felt to be appropriate for the task at hand because while these children are well within the critical period of language acquisition, they are still old enough to produce the amount and variety of data needed for this study as well as to have consistently acquired such related categorical forms as weak past tense and progressive verbs. All studies present a set of challenges to the researcher, and one of those particular to this one is that the sociolinguistic interview techniques used successfully with adults are not appropriate for or very useful with children. Therefore, one of the goals of this study is to modify the sociolinguistic interview for use with very young children by combining its techniques with those which have been proven successful in child language acquisition research.

The variables to be examined were chosen for two reasons. First of all, they are examples of stable variable rules. While there is much to be gained by looking at the contribution of early language learners to the process of language change, it is the aim of this study to evaluate how children acquire patterns of variation which are stable in the adult population. The second reason for the choice of these particular variables is that there is a wealth of literature on both (-t,d) deletion and (ing) production in the adult population.

They are among the most well-documented variable rules in English, and much has been discovered about these rules and their linguistic and extra-linguistic constraints, as will be further discussed in the reviews of literature in Chapters 4 and 5. All of this knowledge is, of course, extremely helpful in comparing the children's acquisition of these variables to that of the adults.

Several questions related to the overall research topic will be presented and discussed. The first of these is, when do the children acquire these two variable rules and their constraints, and how does this time of acquisition compare with that of the learning of related categorical rules? It is hoped that the examination of this question will help to document a closer relationship between the acquisition of variable and categorical rules specifically, and the fields of sociolinguistics and psycholinguistics in general.

Secondly, we will look at some of the explanations for variation that have, at times, been proposed to and often refuted by researchers examining adult patterns of variation. These include the questions of whether or not universal constraints on language or the attachment of probabilities to individual words rather than to abstract grammatical categories can explain documented variational patterns as well as whether the (ing) segments to be studied are underlyingly one morpheme or two for young language learners. It is hoped that this study will add to the knowledge already attained on these issues with the addition of findings from the genesis of these rules in a new generation of children.

1.2 Organization of the chapters

The chapters comprising this work will be organized in the following manner: Chapter 2 will discuss related issues in the literature on child language acquisition, particularly those areas of weak past tense verb formation and the acquisition of progressive verbs. Chapter 3 will present an overview of previous work on child language variation including studies on register variation in very young children. Chapters 4 and 5 will present reviews of the literature on (-t,d) deletion and (ing) production respectively. Chapter 6 will discuss the methodology of the present study, including modifications of the adult sociolinguistic interview process to adapt it for use with children. The findings and discussion of this study as regards (-t,d) deletion will be presented in Chapter 7, while Chapter 8 will present the analysis of (-t,d) deletion as it relates to lexical phonology. Chapter 9 will contain the analysis and discussion of (ing) production. Finally, Chapter 10 will present the summary and conclusions for this study.

CHAPTER 2: RELATED ISSUES IN CHILD LANGUAGE ACQUISITION

2.0 Introduction

The purpose of this chapter is to explore the implications of the acquisition of variability for the study of the acquisition of invariant grammatical forms as pursued in the fields of psycholinguistics and cognitive science. The systems of the past tense and progressive verbal forms will be reviewed and an overview of the psycholinguistic research on these forms given. There will also be a discussion of the implications of the current study for that body of research. Particular attention will be given to the areas of the phonological effects on both variation and past tense acquisition. I conclude by arguing that the importance of evidence from studies of variation in the speech of young children has important insights to offer for our understanding of language learning in general.

2.1 The past tense verb system

The vast majority of English verbs are regular and form their past tense forms by the addition of the *-ed* suffix with phonetic modifications. There are, however, a group of approximately 180 irregular verbs which, while smaller in number, are disproportionately common in the everyday speech of both adults and children. Bybee and Slobin (1982a) divide the irregular verbs into the following classes:

1. No change verbs (e.g. *hit* --> *hit*).

2. Verbs that change final /d/ to /t/ in the past tense (e.g. *build* --> *built*).
3. Verbs that undergo an internal vowel change and add a final /t/ or /d/ (e.g. *lose* --> *lost*).
4. Verbs that undergo an internal vowel change, delete the final consonant, and add a final /t/ or /d/ (e.g. *catch* --> *caught*).
5. Verbs that undergo an internal vowel change whose stems end in a dental (e.g. *bite* --> *bit*).
6. Verbs that undergo a vowel change of /ɪ/ to /ae/ or /^/ (*sing* --> *sang*).
7. Other verbs that undergo an internal vowel change (*give* --> *gave*).
8. All verbs that undergo a vowel change and that end in a diphthongal sequence (e.g. *fly* --> *flew*). (They also include *go* --> *went* in this class.)

2.2 Past tense acquisition

The English past tense verb system can be seen as both more complex and more difficult to learn as compared, for example, to the present progressive system discussed in the following section. Nevertheless, children do learn this system, and this phenomenon has been the focus of many studies in the field of child language acquisition. One of the most well known is that of Roger Brown (1973), who looked at the past tense among other grammatical forms in a longitudinal study of three children, Adam, Eve, and Sarah. Although he concluded that children acquire the irregular past tense verbs before the regular past tense forms, only two of the three children he studied followed that order of acquisition. Brown used as the criterion of acquisition 90% use in obligatory contexts, and, with that as a standard, Adam acquired the irregular past tense form when his mean length of utterance (MLU) was approximately 2.75. Sarah acquired the irregular past tense at an MLU surpassing 2.25, but neither acquired the regular past tense forms until their MLU's were over 4.0. Eve, on the other hand, did not acquire the irregular past tense form until her MLU surpassed 4.0, but acquired the regular form at an MLU of approximately 3.5.

Kuczaj (1976) points out that the criterion of use of 90% in obligatory contexts is problematic in the instance of irregular verbs since errors such as *wented* and *goed* may be considered examples of appropriate and creative use of the regular past tense rule but also are incorrect irregular past tense forms. In fact, in his study using both longitudinal and cross-sectional data, he found that when he looked only at correct past tense uses, such as *went* and *helped*, and errors of omission of past tense, such as *go*, his findings are generally supportive of Brown's. If, however, as in the analysis that Kuczaj prefers, he counted tokens like *goed* as (semantically) correct instances of the regular verb form but also as incorrect instances of the irregular verb form, very different findings emerged. In this case, the children acquired the regular past tense form considerably earlier than they acquired the irregular past tense form.

Whatever methodology for looking at the acquisition of past tense verb forms is used, there appears to be agreement in the literature that irregular past tense forms are the first instances of past marking to appear in children's language (Brown, 1973; Ervin and Miller, 1963). Shortly after the first productions of irregular verbs, regular past tense forms begin to emerge along with overgeneralization errors like *comed* and *ated*. In fact, Slobin (1971) reports instances of overregularizations occurring before correct regular past tense forms. Over time these overregularization errors disappear resulting in a fully acquired past tense system. When this occurs is unclear. Menyuk (1963, 1964, 1969) found irregular past tense errors in first graders. Slobin (personal communication cited in Kuczaj, 1977) reports these errors occurring in 9- and 10-year-old children. Marcus, Ullman, Pinker, Hollander, Rosen, and Xu (1990) estimate a small overregularization rate even in adults of .00004 overregularizations per 1000 sentences of casual conversation[1]. In any event, the rate of overregularization drops precipitously after the pre-school years resulting in the U-shaped behavioral pattern first reported by Ervin and Miller (1963) and later further examined by Bowerman (1982), Pinker and Prince (1988), Marcus et al. (1990), Plunkett and Marchman (1991), Kim, Pinker, Prince, and Prasada (1991), and others. This pattern can be described as a phenomenon in which errors are seen in a behavior which was previously error free. These errors then disappear over time. This tendency to prefer learning by generalizations to learning by individual forms has been discussed both in regard to linguistic development and overall cognitive development (Strauss, 1982).

---

[1]See footnote #34 for further discussion of this estimated overregularization rate.

That overregularization exists is, for the most part, uncontested. Why it exists, as well as why it disappears over time, is the subject of much debate. Early accounts of past tense acquisition were firmly rule based, the assumption being that irregular verbs were learned as exceptions. Berko (1958) showed that children were able to use grammatical rules creatively by adding regular past tense endings to nonsense verbs such as *rik* --> *rikked*. Rumelhart and McClelland (1986), however, provided a connectionist explanation for this process as well as a computer simulation of a verb-learning system which not only generalized regular verb suffixes to new verbs but demonstrated a U-shape learning pattern similar to that of children. They proposed a single learning mechanism in which the input verbs are associated with output units which correspond to the phonological features of a verb's past tense forms. There is no distinction between regular and irregular verbs in this model. Rather, the *-ed* suffix is generalized to new verbs due to the large numbers of input verbs requiring it. This model was further modified by Plunkett and Marchman (1991) who used a multilayer rather than a single layer network and isolated four types of mappings analogous to the relationship between verb stems and past tense forms in English: arbitrary (*go* --> *went*); regular (*hop* --> *hopped*); identity (*hit* --> *hit*); and vowel change (*sing* --> *sang*). The probabilities of association between verbs and past tense types are weighted according to phonological similarity with other stems on which the network has been trained. Plunkett and Marchman also suggest, however, that a child learning this system would have access to semantic information when mapping a stem to its past form.

As previously noted, past tense acquisition was originally conceived of as rule based, and strong arguments for this position continue to be proposed. Kuczaj (1976) put forth two possible explanations for overregularization and its disappearance. First, when the child first produces correct irregular forms, she has analyzed these forms semantically as past but has not syntactically analyzed them as such. He further notes that when the regular past tense rule is learned, it is rarely used to mark an irregular past form as past (e.g. *wented*) but is used to mark present verbs resulting in productions like *goed* and *hopped*. *wented* and others like it, when they occur, are instances of semantic redundancy.

An alternative explanation is that by the time the child learns the past tense rule, she has also begun a syntactic analysis of the irregular past verbs she has been using resulting again in few errors like *wented*. Those that do occur are examples of syntactic, as well as semantic, redundancy (an explanation originally seen in Menyuk, 1963). The child has not, however, attached each past form to its present form, so the productions *goed*, *went*, and *wented* may all co-exist. The matching of these present and past forms is a long and tedious task resulting in the complete acquisition of the past tense or the right-hand branch of the U-shaped curve. Kuczaj further highlights the difference in the acquisition of the regular and irregular verb forms by noting that while MLU is a better predictor of regular past tense acquisition, age is as good a predictor of irregular past tense as MLU. He argues that inferring a grammatical rule is easier than learning many individual forms and may be done with varying amounts of data depending on the child. The learning of individual forms, however, is dependent on exposure, which would increase with the child's age.

Marcus et al. (1990) argue that children have, in fact, united the present and past forms of verbs from an early age and the problem evidenced in overregularizations is not a failure to unify the verb forms but, instead, a failure to retrieve the correct irregular form. When this form is not retrieved, the rule will be utilized resulting in a regular past tense form.

Marcus et al. also note that it is "obligatoriness" which is most related to the onset of overregularization errors. That is, when a child possesses a regular past tense rule and regularly marks verbs as past, these overgeneralizations will appear. Furthermore, these errors are independent of the proportion of regular verbs in either the child's or the parents' speech which, in any case, does not vary greatly. This is a crucial difference between the connectionist and rule based hypotheses since the computer simulation models are dependent upon a large increase in the input of regular verbs to trigger overregularizations.

What is common to all of the above arguments, however, is the observation that while all irregular verbs are open to overregularization errors, not all verbs are overregularized at the same rate. Bybee and Slobin (1982b) note that children tend to overregularize verbs which are used less frequently by their caretakers more often than those used more frequently. They also show that preschool children are less likely to overregularize verbs ending in /t/ or /d/, such as *cut* or *hit* with the exception of the devoicing class of verbs like *build* --> *built*. In addition, verbs which undergo an internal vowel change (and sometimes delete a final consonant) and add a /t/ or /d/ as in *lose* --> *lost*, *say* --> *said*, or *bring* --> *brought* are less likely to be overregularized than those verbs which undergo a vowel change only (*break* --> *broke* and *bite* --> *bit)*. Most likely to be overregularized are those verbs which are characterized by the *ing-ang-ung* vowel change and verbs ending in a diphthong and undergoing a vowel change, such as *blow* --> *blew*.

Bybee and Slobin (1982a) and Marchman (1988) note the presence of what Marchman calls "identity mapping irregularizations." That is, regular verbs are sometimes treated by children as identity verbs (e.g. no phonetic change from present tense to past tense as in *hit* --> *hit*) particularly if they end in a dental consonant. Although they noted that final dental consonants were not necessary to these irregularizations nor sufficient to predict them, Plunkett and Marchman (1990) argue that past tense production is influenced by similarities between stems and past tense forms.

Marcus et al., like Bybee and Slobin, found that "the chief determinant of overregularization-proneness is the verb's frequency in parental speech" (p.57) with high frequency verbs overregularized less often. They also note that there is a small "effect of the verb's phonological neighborhood" and that a cluster of similar irregular verbs can protect a like verb from overregularization but that there is no similar effect for regular verbs.

In summary, the U-shaped past tense learning phenomenon, whether rule based or founded on similarity relationships, is characterized by overregularization errors of irregular verbs which appear after the emergence of the first past tense irregular verbs and disappear over time. While any irregular verb is a candidate for these errors, not all verbs are overregularized at the same rate. Proposed influences on the rate at which a verb is overregularized include its frequency in adult speech, its final consonant, and the similarity between the irregular verb and a group of phonologically similar irregular verbs.

2.3 Acquisition of past tense verbs and variation

There are a number of differences between the study of the acquisition of past tense by children and the study of (-t,d) deletion. The most obvious is that with few exceptions, (-t,d) deletion studies have been restricted to adult speakers. (See Chapter 3 for a discussion of these exceptions.) Secondly, while those interested in past tense acquisition have collected productions of exactly these tokens, variationists, however, do not examine all instances of past tense verbs. Rather, they have collected tokens of words ending in consonant clusters with final /t/ or /d/, which of some have also been instances of past tense verbs. However, even though these pools of data and the conclusions drawn from them are not directly comparable, there is at least one parallel. Researchers in both areas of study have noted the presence of an effect of the phonological form of the item in question and its environment on the likelihood of its undergoing deletion or overregularization. As noted above, according to Bybee and Slobin (1982b), verbs ending in a diphthong were most likely to be overregularized, followed by vowel change verbs ending in a velar nasal, then by vowel change verbs ending in a consonant. Identity verbs ending in /t/ or /d/ were noted by Plunkett and Marchman (1988), Pinker and Prince (1988) and Marcus et al. (1991) to be rarely overregularized.

From the earliest studies on (-t,d) deletion, it has been noted that the presence of consonants in the phonological environment of a final (-t,d) cluster also appeared to have an effect on final stop deletion. Specifically, Labov, Cohen, Robins, and Lewis (1968) found that a non-vowel following a (-t,d) cluster favored deletion more than a following vowel. Studies by Wolfram (1969), Fasold (1972), and Labov (1975) produced similar findings. These same researchers also noted that the tendency to delete /t/ or /d/ was also affected by the segment preceding the cluster. Labov et al., for example, found that preceding obstruents favored deletion more than preceding sonorants. Finally, Labov (1989) found that both a seven-year-old and his parents deleted final /t/ and /d/ more often when the cluster was preceded by a third consonant than when it was preceded by a vowel.[2]

Just as factors such as grammatical form, syllable stress, and others affect (-t,d) deletion, other factors are influential in determining overregularization as well. For example, it was previously noted that a verb's frequency in parental speech affects its likelihood of being overregularized as does whether or not an irregular verb is a member of a family of phonologically similar verbs. It appears likely that a fruitful avenue of future research would be to apply variable rule analysis methodology to the

---

[2]See Chapter 4, Sections 4.1.1.1 and 4.1.1.2 for a more detailed discussion of these constraints on (-t,d) deletion.

overregularization phenomenon in young children in an effort to quantify the effects of the factors mentioned above. Of particular interest would be to note whether or not there are similar phonological factors influencing overregularization and (-t,d) deletion.

2.4 Overview of the Progressive Verb

The {ing} morpheme can be divided into four grammatical categories.

Progressive participle. *She is singing a song.*

Present participle. *She kept singing the song.*

Gerund. *Her continuous singing of that song drove everyone crazy.*

Verbal adjective. *Her singing voice left much to be desired.*

The following summary of the acquisition of this system will be restricted to the progressive participle because this is the form most successfully acquired by preschool children although a limited number of examples from the other categories will be presented in the results of this study. (Please see Chapter 5 for a more thorough discussion of the (ing) variable.)

2.5 Acquisition of the progressive participle

There is far less controversy over the acquisition of the progressive participle than over the past tense verb. This is perhaps not coincidental to the finding that the progressive verbs are also easier for children to learn. Brown (1973) and deVilliers and deVilliers (1985) note that the {ing} morpheme is among the earliest grammatical forms acquired -- emerging before or in conjunction with the plural noun inflection. Acquisition, in this case, however, refers only to the {ing} morpheme itself; the *be* auxiliary emerges afterward. In fact, Cazden (1968) notes that none of the three children she (and Brown) studied had reached the 90% criterion for the auxiliary by the end of their longitudinal study at which point the children's MLU's had topped 4.0. The -*ing* suffix was acquired when the children's MLU's were approximately 2.25. Brown also noted that unlike in the cases of the plural, regular past tense, possessive, and third person singular present inflections, the progressive verbs were never overgeneralized. Brown suggested, however, that children do not learn the rule that process verbs, as opposed to state verbs, may form progressives. Rather, children simply learn that some verbs are "-*ing*-able", and some are not. His suggestion is reinforced by the close parallel he found between the children's and their parents' use of the progressive form with the same restricted set of verbs.

Kuczaj (1976), whose fourteen cross-sectional subjects and one longitudinal subject had MLU's in excess of 2.5, found only one possible overgeneralization of the progressive. He did, however, find several novel uses of -*ing*, including the following:

1. *It's weathering out there, too. Why is it weathering?* (it is storming) (H. K., 3;6) p.31.

2. *I'm shirting my man.* (putting a shirt on a doll) (J. W., 5;7) p.32.

Kuczaj distinguishes the above examples from overgeneralizations in that they are appropriate uses of the progressive on "novel" verbs, not violations of the process-state distinction as would be *knowing* or *having* (possessive). Kuczaj concludes that children begin with an -*ing*-able/non-*ing*-able distinction but later learn or discover the process-state distinction, which allows for these creative uses of the -*ing* inflection.

Studies on the acquisition of variation overlap the interests of those who study the acquisition of invariant grammar with respect to the nature of learning in general as well as to parallels between findings on specific grammatical forms. As discussed above, there is disagreement as to whether children acquiring language are forming associative networks with weighted connections or learning rules. The associated input and output units of the single learning mechanism put forward by Rumelhart and McClelland and Marchman and Plunkett are based on phonological similarity and, possibly, semantic information, not grammatical category. In fact, Kim, Pinker, Prince, and Prasada (1991) suggest that the connectionst approach to language learning makes not only rules but grammatical categories obsolete in that this approach implies that children may not have mental representations of rules, verb roots, or lexical items. Rather, past tense formation would be accomplished as a unified process of phonological association whether the verb itself is regular or irregular. Kim et al. go on to refute this claim by demonstrating that adults do rely on grammatical categories in formulating past tense forms of novel verbs. They found that verbs with noun roots -- denominal verbs -- were judged better with a regular past tense (e.g. *line-drived* as in the past of "to hit a line drive" was preferred to *line-drove*.) On the other hand, subjects preferred irregular past tense forms for verbs with verb roots (e.g. *line-drove* was preferred for the past of "to drive along a line".) Previous studies on (-t,d) deletion and (ing) production by adults also demonstrate the need for grammatical categories as well as rules in adult grammars since speakers, in following the grammatical constraints for both of these variables, show a sensitivity to grammatical categories that belies their obsolescence. A previous study on (-t,d) deletion (Guy and Boyd, 1990) suggests that young children show this same sensitivity. (See Chapter 4 for further discussion of this and other related studies.) The current study seeks to discover whether children near the beginning of the language acquisition period also share this sensitivity. If so, then the argument for rule learning based on grammatical categories is strengthened, particularly since, at three and four years of age, these children would be learning these rules before any direct teaching of grammatical categories would take place. Therefore, the study of the acquisition of variation appears to be an important corollary to that of the acquisition of categorical rules in the search for an overall model of child language acquisition.

CHAPTER 3: PREVIOUS STUDIES IN CHILD LANGUAGE VARIATION

3.0 Introduction

This chapter will provide an overview of research on the subject of the acquisition of variation. Although a number of studies have looked at linguistic variability in older elementary school and adolescent children (Labov, 1966, 1972; Reid, 1978, for example), few have focused on preschool and early school aged children. This chapter will review these studies which examine variation in the very early years.

3.1 Acquisition of register

While studies on variation in children have been rare, there have been many studies on a related topic -- the acquisition of register. These studies of register in adults and children are sometimes grouped with those of stylistic and social variation, but there are differences as well as similarities between them. Overlapping vocabulary in the two areas of study may contribute to the frequent lack of distinction between them. The word *style*, for example, is used in different ways in both areas of research. It can mean a register, a level of formality, or an informal assessment by a researcher of a speaker's communication. The word *sociolinguistic* as well has been used differently at times. For example, the sociolinguistic skills discussed by Andersen (1986, 1990) are primarily those of register, not those having to do with variable rules or social variation. On the other hand, register is rarely mentioned at all in variational research. While it might be agreed that both areas rightly belong in a comprehensive discussion of sociolinguistic skills, there also seems to be value in distinguishing between them for the sake of clarity in discussion. Andersen distinguishes between dialects which vary according to the *users* of language and registers which vary according to the *use* of language. This is a useful distinction in that it does differentiate between register and socially or regionally based dialects, but it does not account for what are usually referred to as stylistic constraints on variation.

One distinguishing feature found in many studies of register which may be useful in differentiating register and what I will call variational style is the concept of role. Role is often dependent on the relative status of the speakers (e.g. ages, professions, sexes, kinship roles, or native languages) as well as the topic and setting of the interaction. A person may "play" many roles in the course of a lifetime, or even a day, but each role is seen as a discrete entity. Stylistic variation, on the other hand, is often conceived of as a continuum of formality. One of the most common demarcations of variational style is based on tasks engaged in within the interview setting, such as the narrative, informal conversation, reading passages, and word lists. Even variation according to the addressee, which is usually included in discussions of variational style, is generally seen as a condition affecting the formality of the interaction rather than a set of role demarcations. In fact, as discussed in chapter 4, Bell (1977) found addressee to be a primary determinant of style in his study of radio news readers in Auckland, New Zealand.

Another way to distinguish between register and style is to view register as a precursor to style -- a proto-style[3] -- or simplified version of what will later be a speaker's full stylistic range. For example, many of children's early register changes involve the simplification of language, as in baby talk. Even when changing to a "father register", for example, the result is an increase in simple, direct imperatives, as well as the lowering of the voice. It is possible that children's range of register in the preschool years may be confined to simplifying their speech and changing suprasegmental features, such as pitch and loudness.

Research indicates that very young children have begun the process of acquiring different registers. Baby talk (Ferguson, 1977) is a simplified register which has been found to characterize parents' speech to infants in many of the world's societies. It includes grammatically and phonologically simplified utterances, exaggerated intonational contours, higher pitch, and increased use of diminutives. It is a register most babies and young children often hear, and it is also one they have been heard to produce at an early age. Shatz and Gelman (1973) found that 4-year-olds tailored their instructions to meet the perceived needs of 2-year-old listeners. They used some of the features of baby talk and were, at times, even able to state their assumptions of the younger children's communicative needs. These findings that young middle class children have the rudiments of a baby talk register have been replicated with working class children (Miller and Garvey, 1984) and children from other cultures (Watson-Gegeo and Gegeo, 1986). Ervin-Tripp (1973) noted that when

---

[3] I am indebted to Gregory Guy for suggesting the concept of proto-style.

preschool children role-played, they demonstrated consistent speech patterns depending on the role they were playing -- mother, baby, doctor, etc. Andersen (1986, 1990) recorded children aged 3:9 to 7:1 engaged in semi-structured role play situations using puppets. She found that by the age of 5, children could "choose sentence structures, lexical items, and phonological features to fit the different roles in their sociolinguistic repertoires" (page 159). For example, when children role played a father talking to a child, they adopted a deeper voice and backed and lowered vowels, and used a large number of direct imperatives. As "mothers" the children used a higher pitch, softer voice, more endearments and babytalk terms, and fewer imperatives.

As the discussion to follow will elaborate, the acquisition of variational style is to a large extent unexplored, and the overall findings are not clearcut. While there are studies which show older children do demonstrate stylistic variation (See discussion of Fischer, 1958 and Romaine, 1978 below.), the early assumption that children are monostylistic until sometime around adolescence (Labov, 1970; Wolfram and Fasold, 1974) has gone largely unchallenged until recently. The results of Purcell, 1984, who found indications of stylistic variation in her 5- to 12-year old speakers but did not break down her findings on stylistic variation by age, and Labov (1989a), who found it in a 6- but not a 4-year-old, suggest, however, that stylistic variation is present earlier. Current indications are that the stylistic variation that preschoolers may be confined to proto-stylistic or register changes. Further research is required to determine when the emergence of variational style occurs.

3.2 Acquisition of variable rules

The acquisition of variable rules by children and the situations which encourage or discourage this variation is an area of study that sociolinguists are just beginning to explore. With few exceptions, these studies are of stylistic and social variation, and few of them include preschool children. The first study in this area was that of Fischer (1958) who found social variation of (ing) in children aged 3 to 10 and stylistic variation in a 10-year-old. It is notable that Fischer included 3- and 4-year-olds in his study. He did not, however, separate the data by the ages of the children, so it is impossible to know to what extent the preschool children had an impact on his overall findings. He found that girls and "model" boys -- that is, those who were judged by their teachers to be especially well-behaved -- used more of the [iN] form than the "normal" boys who used more [ɪn]. He also noted that the 10-year-old boy whose speech he analyzed used [iN] almost exclusively in a formal testing situation but used mostly [ɪn] in a less formal interview.

Romaine (1978) continued the exploration of social and stylistic variation in children by looking at the production of word final /r/ in Scottish English by 6-, 8-, and 10-year-old boys and girls and found sex, age, and some style variation. She noted that the girls in her study used more of the [r] variant, while the boys were more often r-less. She concluded that the girls were participating in a change from above the level of consciousness and favoring a prestige variant, and the boys were participating in a change from below the level of consciousness and favoring a variant with less, or perhaps covert, prestige.

Purcell (1984) studied social and style variation operating on several variables produced by 5- to 12-year-old speakers of Hawaiian and "General" American English. She found that ethnicity and social class were more important than age in determining shifts from General to Hawaiian English and vice versa. She also noted that these shifts often accompanied changes in style factors such as addressee, topic, emotion, or genre shift. She, like Fischer, did not break her findings down by age, making it impossible to determine the contribution of her youngest subjects. Nevertheless, her results are encouraging in showing the sensitivity to stylistic and social factors in the pre-adolescent years.

One of the first studies to look at variation in preschool children was also one of the first to examine the linguistic as well as the social constraints on that variation. Kovac and Adamson (1981), in their study of finite *be* in African American and white English speakers, considered the question of developmental versus dialectal variability by recording 3-, 5-, and 7-year-old African American and white children from middle and working socioeconomic class backgrounds. Their results revealed that for the white children of both social classes absence of finite *be* appeared to be developmental in nature unless clear evidence of peer influence was present. For the African American children, however, the results varied by socioeconomic group. African American working class children showed an increase in deletion between ages 3 and 5, followed by a decrease in deletion along with an increase in contractions and full forms between ages 5 and 7. On the other hand, their middle class peers exhibited a decrease in deletion (and an increase in contraction) between 3 and 5 followed by an increase in deletion (and decrease in contraction) between ages 5 and 7. Full forms remain constant in both groups. In other words, working class African American children learned deletion before middle class children, while contraction appeared to precede deletion for the middle class children. The authors suggested that an examination of peer influence might help to explain the upsurge of deletion in the 7-year-old middle class African American speakers. The constraints on the rule of deletion were even more difficult to acquire than the rule itself. Although both the grammatical and phonological constraints for contraction had been acquired by both groups of African American children by age 3, the constraints on deletion typical of adults had not been completely learned even by age 7.

Guy and Boyd (1990) looked at only the grammatical constraints of (-t,d) deletion in their study of its acquisition by speakers aged 4 to 65 in "semi-weak" or "ambiguous" past tense English verbs such as *lost, sold,* and *slept*. They proposed a three stage course of development for this class of verbs. For young children, the semi-weak verb class did not appear to exist in that the final *t*'s and *d*'s were categorically omitted. In Stage 2, the verbs appeared to be analyzed by the speakers much like monomorphemic words, for which the speakers demonstrated a similar probability of deletion. The semi-weak verbs were not formulated as a separate morphological class until Stage 3 -- a stage not attained by all speakers even in adulthood. These results suggest that language acquisition continues into adulthood, long after it has usually been assumed to be complete.

Finally, Labov (1989a) studied stylistic and linguistic variation in children in the King of Prussia area of Philadelphia. He found that a 7-year-old replicated his parents' pattern of (-t,d) deletion in regard to stylistic and stress variation. This speaker also demonstrated phonological and grammatical conditioning similar, although not identical, to that of his parents. In fact, the primary difference was that the boy deleted (-t,d) in verbs in the semi-weak class at a similar rate to those in the monomorphemic class, whereas his parents showed deletion in the semi-weak verbs similar to that in weak past tense verbs (e.g. *walked* and *jumped*.) These results corroborate those of Guy and Boyd discussed above. In the same study, an examination of (ing) variation also indicated that children were beginning to master the constraints on this rule. Labov found that the 7-year-old had learned both the grammatical and stylistic patterns of (ing) variation, and a 6-year-old had acquired the stylistic but not the grammatical constraints. A look at limited data on a 4-year-old revealed no indications that he had mastered the constraints on (ing) production at all.

One of the questions proposed by Labov (1989a) is that of whether children first acquire intralinguistic or extralinguistic constraints on variation. Unfortunately, it is very difficult to answer this question since almost all of the studies discussed above focus on one type of constraint or the other, but not both. It is clear from the results of Fischer, Romaine, Purcell, Kovac and Adamson, and Labov that social variation, whether it be social class, age, or sex, is present at however early an age researchers have looked for it. There are fewer findings on stylistic variation although early indications are that it may be taking place as early as age 5 or 6. Studies of register have noted the presence of register variation very early in the pre-school years. Romaine found evidence of style shifting in 6-year-olds, and Purcell may have found it in 5-year-olds. On the other hand, both Kovac and Adamson and Guy and Boyd found indications that the acquisition of grammatical variation was also taking place at 3 or 4 years of age. Labov, the only one to look at both stylistic and linguistic constraints on variation in one 6-year-old found evidence that stylistic constraints are acquired first, but clearly more studies are needed to formulate a more definitive answer.

3.3 Conclusions

As the previous discussion has shown, there have been several studies beginning to explore the acquisition of variability. Until recently, most sociolinguistic studies of very young children examined register, not stylistic, social, or linguistic variation. On the other hand, most variational studies have concentrated on speakers who were of adolescent age or older. The variational studies presented above were unique in including preadolescent children as subjects, but only the work of Kovac and Adamson, Guy and Boyd, and Labov have included children under the age of 5, and only Kovac and Adamson have looked at 3-year-olds. As psycholinguistic research has shown, much of the work of the acquisition of categorical rules takes place in the preschool years, and most grammatical forms are already in place by age 5. As will be stated repeatedly in the course of this study, there remains a need for studies of linguistic and extralinguistic variation in very young children to fill this gap and contribute to the knowledge of how these rules and constraints, so well-documented in the adult population, are passed from generation to generation.

CHAPTER 4: AN OVERVIEW OF THE VARIABLE (-t,d)

4.0 Introduction

This chapter will present an overview of the sociolinguistic literature on (-t,d) deletion with particular emphasis on the numerous studies of this phenomenon in the adult population. Chapter 7 will present the findings of the present study on (-t,d) deletion in children and compare the results with those on adults.

The domain of (-t,d) deletion can be seen as a subset of those consonant clusters which are reduced by a process of simplification. Other examples of this process include [l] vocalization (e.g. [kold --> ko:d], nasalization (e.g. [want --> wa$t], and reduction of final consonant clusters ending in other stops (e.g. [dEsk --> dEs] or [lɪsp --> lɪs]. (-t,d) deletion results when the final alveolar stop is omitted from a word final consonant cluster (e.g. last night --> [lœs nait] or missed you --> [mɪs yu].) (-t,d) deletion has been a frequently studied variable for sociolinguists because the final /t/ and /d/ clusters occur frequently in the English language, and their reduction is widespread. In addition, the contrast between the (-t,d) monomorphemic final cluster, as in *mist* and the *-ed* past tense suffix, as in *missed*, allows for an examination of morphological as well as phonological constraints on (-t,d) deletion. These two types of constraints will be the focus of most of the rest of this chapter along with stylistic and social factors which affect deletion. A discussion of some recent work on (-t,d) deletion and lexical phonology will be presented in Chapter 8.

4.1 Internal constraints

The internal constraints on (-t,d) deletion include phonological, grammatical, and prosodic constraints all of which will be discussed in this section.

4.1.1 Phonological constraints on (-t,d) deletion

Following segment and preceding segment are the two phonological factors affecting (-t,d) deletion.

4.1.1.1 Following segment

In 1968, Labov, Cohen, Robins, and Lewis provided a detailed look at (-t,d) deletion which included the effect of the following segment constraint. For their study, following segment was divided into vowel vs. nonvowel, and a following nonvowel was found to favor deletion. Wolfram (1969) found a similar ordering of constraints on deletion when he divided the segments into consonants vs. nonconsonants as did Fasold (1972) and Labov (1975) when they opposed consonants and vowels.

Phonemes can be described by their distinctive features, and, as such, consonants are [+cons,-voc]. Vowels, on the other hand, are described as [-cons,+voc]. If consonants promote deletion and vowels inhibit deletion, then, as Guy (1980) pointed out, liquids, [+cons,+voc], and glides, [-cons,-voc], could be expected to be somewhere in the middle. This was found to be the case by Labov et al. (1968) and Labov (1975) who reported glides to favor deletion more than vowels but less than consonants. Neu (1980) ran a chi square analysis on the following segment constraint on (-t,d) deletion on 15 white speakers from various geographical areas and found somewhat different results in that liquid and glide did not behave significantly differently from consonant as a constraint on (-t,d) deletion. All three, however, favored deletion significantly more often than vowels. Guy (1980) was in agreement with the Labov findings and reported the following order of deletion: consonant > liquid > glide > vowel.[4]

A word-final consonant cluster can be followed by a pause and/or a sentence boundary as well as a phonological segment, and the effect of following pause has been found to be less consistent in its order than the following segments. Labov et al. (1968) grouped pause with the consonants in their study of African American and white speakers in New York City. On the other hand, Wolfram (1969), when examining the speech of Detroit African American and white speakers grouped following pause with the nonconsonantal segments. Fasold (1972) divided his data into following consonant, vowel, and pause in his study of (-t,d) deletion in African American Washingtonians and found that following pause had an effect on deletion similar to that of consonant. Labov (1975) found that for eight white subjects, pause and vowel had similar effects on deletion. Wolfram and Christian (1976) noted results similar to those of Labov in their study of six white Appalachian speakers.

To shed light on these apparently different analyses of following pause, Guy (1980) examined (-t,d) deletion in Philadelphia and New York speakers. He found that for New Yorkers, following pause is a favorable environment for deletion, while for

---

[4]The symbol > is to be read as "favors deletion more than."

Philadelphians, it is an unfavorable one. He further notes that these dialectal differences in the effect of following pause may explain why Labov et al. treated pause as equivalent to consonant, while Wolfram treated it as equivalent to vowel. Guy states that although white speakers appear to vary by geographic region in their analysis of following pause, there is some evidence that African American speakers all show an effect of following pause favorable to deletion (Fasold, 1972; Labov et al., 1968). Finally, Guy points out that this variability in the effect of following pause confirms Labov's suggestion that the consonant > liquid > glide > vowel constraint ordering may be universal and functionally based, whereas the relative effect of following pause is dialectally based and not universal.

4.1.1.2 Preceding segment

Guy (1980) points out that in English there is a tendency toward anticipatory assimilation, and, therefore, following segment usually has a greater impact on variable rules than preceding segment. The effect of preceding segment, however, cannot be, and has not been, discounted. Labov et al (1968), for example, found an obstruent > sonorant hierarchy for both African American and white New York City speakers.

Wolfram (1969) and Fasold (1972) looked at the effect of preceding segment only in bimorphemic clusters, and, therefore, their results are not directly comparable to the other studies discussed in this section. Their results also differ from each other. Wolfram found a stop > spirant > sonorant ordering for the African

American and white Detroit speakers he studied. In contrast, Fasold found that sonorants most favored (-t,d) deletion in Washington African American speakers, followed by spirants and stops.

Neu (1980) found the following order of the effect of preceding segment for the men in her study: sibilant > nasal or stop > liquid or fricative. (The women whose speech she examined showed basically the same order except that for them the effect of preceding sibilant did not differ significantly from that of nasals and stops.) Guy (1980) and Neu used the same categories of preceding segment, and Guy's results were as follows: stop > sibilant > nasal > fricative > liquid. He noted, however, that the placement of fricative in the hierarchy is problematic since there is disagreement between the individual and group analyses as to its placement. There is also a reversal of sibilant and nasal in one of his data sets.

In a study of (t) deletion in a Dutch dialect, van Hout (1989) reported similar results. He found that there is a weakening of final (t) corresponding to its difference in sonority with the preceding consonant. Therefore, the final (t) is strong after nasals and liquids (as well as vowels and glides.) On the other hand, (t) is more likely to be reduced or deleted after an obstruent (fricative and plosive.)

In general, there is less overall agreement in the results of the analyses of preceding segment than in that of following segment, except for following pause. Unlike following pause, however, there is no evidence to suggest dialect differences influencing the results on preceding segment. Two possible reasons for the relative lack of agreement in the preceding segment findings are that there is a comparative lack of data for some of the preceding segment environments and that the overall effect of preceding segment is less than that of following segment. It must be noted, however, that while the latter possibility may have some merit for English, van Hout notes that in Dutch dialects, preceding consonant is the most powerful factor influencing deletion.

4.1.2 Grammatical constraints on (-t,d) deletion

The most consistently found effect of grammatical form is that of monomorphemic versus bimorphemic words. Words containing only one morpheme, such as *mist*, are more likely to undergo (-t,d) deletion than words containing two morphemes, such as *missed*. These two-morpheme words are regular weak past tense verbs whose final alveolar stop suffices make up the past marker. There are several other ways of marking past tense in English, however, one of which is also relevant to (-t,d) deletion. (See Chapter 2 for a more detailed discussion of English past tense.) Verbs such as *leave*, *sell*, and *sleep* form their past tense forms by means of a vowel change and the addition of an alveolar stop (and in the case of verbs like *leave* devoicing the present tense final consonant) resulting in the final consonant clusters *left, sold*, and *slept*. These verbs were labeled ambiguous verbs (Labov et al., 1968) in that they contain a double marking -- both a vowel change and an -*ed* suffix. Unlike the case of the weak past tense verbs, omission of the final stop in ambiguous verbs does not remove all past marking of the verb. These verbs have also been called semi-weak verbs to distinguish them from weak past tense verbs such as *call/called* and strong verbs such as *fight/fought*. Chomsky and Halle (1968) describe these semi-weak past tense verbs as having a different morpheme boundary (+ as in /slEp+t/) from weak past tense verbs which contain a standard morpheme boundary (# as in /kAl#d/). The question becomes one of whether speakers treat these verbs like monomorphemes, possibly associating them with strong ablauting verbs like *eat/ate*, like weak past tense verbs, or like a separate and distinct class of words.

The tendency for monomorphemic verbs to promote (-t,d) deletion and for weak past tense verbs to disfavor it is almost uncontested in the sociolinguistic literature in English regardless of geographic or social dialect. It has been found to be true for African American Vernacular English (AAVE) speakers in New York City (Labov et al., 1968; Wolfram, 1972) and Washington D.C. (Fasold, 1972). White speakers from Appalachia (Wolfram and Christian, 1976), New York City (Labov et al, 1968), and Philadelphia (Guy, 1980) have also been found to favor (-t,d) deletion in monomorphemes and disfavor it in weak past tense verbs. Wolfram found the same results in speakers of Puerto Rican English. In addition, three researchers looked at the geographical consistency of the grammatical constraint by examining the speech of residents of several areas of the United States and, in one case, England. Labov (1975) found consistency in a number of American and British dialect areas. Guy (1980) found this pattern in his study which included Philadelphians, New Yorkers, and a group of "cosmopolitans" from various dialect regions of the United States. Neu (1980) looked at speakers from California, Ohio, Michigan, Baltimore, Nebraska, Missouri, Massachusetts, and New York City and, again, found monomorphemes to favor (-t,d) deletion and weak past tense verbs to disfavor it. The one exception in this literature was noted by Bell (1977) who found no effect of a morpheme boundary in his study of consonant cluster reduction in Auckland radio news style English. He stated that one reason for the lack of significant findings in this area may have been the formality of the news reading style he studied.

Several possibilities as to the reason for the strength and consistency of the effect of grammatical form have been proposed. One is that it may have to do with a resistance to delete semantically salient morphemes such as the *-ed* in *missed* as compared with the /t/ in *mist* which does not convey semantic information. The presence or lack of a morpheme boundary has also been suggested as a contributing factor. A third explanation based on lexical phonology has been proposed by Guy (1991) and will be discussed in Chapter 8.

The position of the semi-weak verbs in this hierarchy was not entirely clear at the outset. Labov (1975) and Guy (1980) found that the semi-weak verb was more favorable to (-t,d) deletion than the weak past tense form but less favorable than the monomorpheme (i.e. monomorpheme > semi-weak > weak past tense) in their studies of white speakers. Labov et al. (1968) also found this constraint ordering in their analysis of (-t,d) deletion in African American speakers. (The other studies of speakers of African American Vernacular English did not include all three grammatical forms in their analyses.) These results suggest an analysis of these verbs as a separate category from both monomorphemes and weak past tense -- one in which the past tense /t/ or /d/ is preserved to a greater extent than it would be were it a part of the preceding morpheme but not as much as in the cases where (-t,d) is the only indicator of past tense. Neu (1980), however, found the semi-weak verb class to be not significantly different from the weak past tense verb class in its (-t,d) deletion rate.

Guy (1980) suggested that there may be social factors influencing the (-t,d) deletion rate in the semi-weak verb class. He found that among his working class speakers, the deletion rate for semi-weak verbs was as high as, if not higher than, the monomorphemes. His middle class speakers, however, deleted (-t,d) at a rate between that of monomorphemes and weak past tense verbs. He further noted that children (in this study, aged 8, 9, 10, and 14) also showed high deletion rates for semi-weak verbs leading to the conclusion that the high A-lect, as he called this pattern, might be a developmental stage in language acquisition. He further expanded on this idea in his study with Sally Boyd (1990) which was described in more detail in Chapter 3. To review briefly, they proposed a 3-stage course of acquisition. In Stage 1, (-t,d) deletion was categorical in semi-weak verbs; thus the alveolar stops did not appear to exist in these verbs for young children. In Stage 2, semi-weak verbs did not differ significantly in their deletion rate from monomorphemes; and in Stage 3, the semi-weak verbs were analyzed as a separate morphological class and were deleted at a rate between monomorphemes and weak past tense verbs. Labov's (1989a) results were supportive of this course of acquisition, at least insofar as its last two stages were concerned, in that he found that a 7-year-old deleted (-t,d) at a rate not significantly different from monomorphemes while his parents showed the monomorpheme > semi-weak > weak past ordering.

4.1.3 Relative weighting of grammatical and phonological constraints

One of the questions emerging even from the earliest of the (-t,d) deletion studies was whether the grammatical or following segment constraint had the stronger impact on the application of the deletion rule. Labov et al. (1968) looked at the cases where one effect is favorable to deletion (i.e. following consonant) and one effect is not favorable to deletion (i.e. weak past tense verbs). They found that the answer to this question varied depending on the speech community examined. For almost all of the lower and working class peer group speakers of African American Vernacular English, they found the following segment constraint to be stronger. For the white New Yorkers and middle class African American speakers, however, the two constraints were equally strong. They noted, however, that the effect of the grammatical form constraint grew stronger with increased age and socioeconomic class for the African American speakers. The upper section of the working class adults raised in the north, as well as the middle class adults, reversed this pattern in careful speech showing the grammatical constraint to be the stronger one. Only two of the six younger groups showed even an equal weighting of constraints, and one of these was the Lames who were notable in not belonging to the dominant peer group in their area.

Fasold (1972) agreed with the findings of Labov et al. He looked at (-t,d) deletion in weak past tense verbs by Washington D.C. African American speakers. Following segment was found to be the major constraint followed by syllable stress and grammatical category. Neu (1980) used a chi square analysis of her data and also found the grammatical form constraint to be the stronger in the cross dialectal group of white speakers she studied.

Summerlin (1972) examined (-t,d) deletion in rural Southern African American and white lower socioeconomic class second graders and high school students and African American and white teachers. She also compared her results to those of Labov et al. (1968). She found no difference in deletion rates between second graders and high school students although African American speakers deleted (-t,d) more than white speakers. In contrast to the findings of Labov et al. and Fasold, African American second graders and high school students did not show the phonological constraint at all although the African American teachers and white speakers of all ages did. All groups demonstrated the grammatical constraint. Summerlin studied several variables along with (-t,d) deletion including post-vocalic /r/ and /l/ and word-initial and word-final /Q/. She found that overall, with the exception of the phonological constraint on (-t,d) deletion, Northern and Southern African American speakers spoke similar dialects, while African American and white Southern speakers spoke different dialects.

Wolfram (1972) studied (-t,d) deletion in Puerto Rican and African American teenage male subjects. He found that the constraint ordering for both groups of subjects was following segment > syllable stress > grammatical constraint. These findings are in agreement with those of Labov et al. and Fasold (1972) and indicate that in terms of constraint ordering, AAVE and Puerto Rican English are similar.

Two researchers found exceptions to the generalization that white speakers show a stronger effect of grammatical constraint than phonological constraint. Wolfram and Christian (1976) found the phonological constraint to be the strongest in their study of white Appalachian speakers. Bell (1977) found that the Auckland news readers he studied showed no effect of grammatical form but did demonstrate an effect of the phonological constraint on consonant cluster reduction.

In summary, most of the research shows that African American and white speakers show different orders of importance of the constraints on (-t,d) deletion. For African American and Puerto Rican speakers phonological constraints are more important than grammatical constraints. (This generalization is more accurate for younger African American speakers since, as Labov et al. (1968) point out, the effect of the morphemic boundary increases with age in these speakers.) For white speakers, it is the grammatical constraint which is more important. The exceptions are the Southern African American younger speakers studied by Summerlin who did not show the phonological constraint at all, the white Auckland speakers studied by Bell who did not show the grammatical constraint, and the white Appalachian speakers examined by Wolfram and Christian who showed the same order as the majority of the African American speakers of phonological constraint over grammatical constraint.

4.1.4 The effect of a prosodic factor: Syllable stress

Few of the studies on (-t,d) deletion looked at syllable stress. Those that did, however, (Fasold, 1972; Wolfram, 1972) found unstressed syllables to favor deletion more than stressed syllables. Labov (1989a), in his study of deletion in adults and children, found that both a 7-year-old boy and his parents deleted (-t,d) more often in unstressed than in stressed syllables. These results are not surprising in that deletion is a linition process which could be expected to take place more frequently in unstressed syllables.

The size of the effect of stress on (-t,d) deletion appears to be relatively small. Fasold, for example, labels it a "fourth-level constraint." Guy (1980) also found the strength of the effect of stress to be minor although he notes that these differences in syllable stress "may account for a small portion" of the individual differences that he found.

4.2 External constraints

The external constraints on (-t,d) deletion covered by this review will include stylistic factors and the social factors of race, age, sex, and socioeconomic class.

4.2.1 The effect of style on (-t,d) deletion

Stylistic variation is often related to social class variation in sociolinguistic studies. That is, as interactional style becomes more formal, there is often a shift toward the variant most used by middle class speakers, which is also the prestige variant. One would expect to find, then, that as informality increases, so does deletion of (-t,d);

and this is exactly what has been found to occur. Labov (1967) and Wolfram (1969) both found a more informal speaking style to promote (-t,d) deletion more than a formal style. Labov (1967) notes, in addition, that style can be a variable of considerable importance. He recorded a African American woman in both an informal "family interaction" style and in a more formal interview style. He found that only in the more careful style did she differentiate by means of deletion the monomorphemic words from the regular past tense verbs.

Guy (1980) points out that the methods for eliciting a variety of styles during the sociolinguistic interview results in comparatively small amounts of data for each style. Guy did, however, look at style shifting in two white subjects who showed clear style differences during the interview. Both of these speakers showed increased probabilities of (-t,d) deletion during informal portions of the interview and decreased deletion during the more formal portions. His findings also revealed that style tends to be associated with a general upward or downward shift in factor weights but does not change the relation of the probabilities to each other, a finding that is in contrast to those of Labov on African American speakers.

Bell (1977) proposed a different analysis of style in his study of news readers in Auckland, New Zealand. He stated that a speaker changes his or her speaking style toward the dialect of the addressee. His study of consonant cluster simplification, among other variables, showed that cluster reduction decreases with the increased social status of the radio audience of a particular station. Even when the same news reader spoke on two different radio stations, his speaking style shifted to mirror that of the audience of that station. In other words, news style varies primarily with the audience variables and only secondarily with the speaker characteristics.

4.2.2 Social factors

4.2.2.1 The effect of race on (-t,d) deletion

One reason that (-t,d) deletion is so interesting to sociolinguists is precisely the fact that it is so consistent across social and geographic lines. It is therefore unusual to find large differences in the effects of the various linguistic constraints on this rule. The differences, with few exceptions, tend to be those of overall rate of deletion and the strength of the various constraints. Summerlin (1972), for example, found that her African American subjects tended to delete final stops more often than her white subjects. Wolfram (1969) reported similar findings but noted that there appeared to be an even stronger effect of racial isolation. That is, in his upper middle class subjects, the white speakers and the African American speakers who had a lot of contact with whites in their daily lives showed similar low rates of deletion. The African American speakers with predominantly African American contacts, however, showed a higher rate of deletion.

An exception to the generalization that African American and white speakers show similar responses to the various constraints on (-t,d) deletion is the case of following pause. As was previously discussed, following pause varies in its effects according to geographical area. New Yorkers, for example, evidence increased deletion with following pause whereas Philadelphians show decreased rates of deletion. The studies of Fasold (1972), Labov et al. (1968), and Guy (1980) all indicate that African American speakers speak a dialect evidencing high probabilities of (-t,d) deletion before pause regardless of geographical area.

The differences in the relative strengths of the phonological and grammatical constraints on (-t,d) deletion have been discussed previously. Please see Section 4.1.3 for this discussion.

In summary, the differences between African American and white speakers in terms of (-t,d) deletion are that African American speakers delete (-t,d) at a higher rate than white white speakers. African American speakers also show high rates of deletion following a pause, whereas whites vary in the effect of following pause depending on geographical dialect. These results are consistent with other studies (Labov, 1966, for example) showing that African American speakers, like men and working class speakers produce more of the less prestigious variants than women and middle class speakers.

4.2.2.2 The effect of socioeconomic status on (-t,d) deletion

Results of sociolinguistic studies on a number of variables (Labov, 1966; Trudgill, 1974) indicate that the higher the socioeconomic status of the speaker, the more likely it is that he or she will use a prestige variant. The findings for (-t,d) deletion are no exception. Wolfram (1969) examined the speech of upper middle, lower middle, upper working, and lower working class African American speakers and found that the lower the socioeconomic class, the greater the probability of (-t,d) deletion. In addition, the non-consonantal following segment provided a more diagnostic environment for these social class differences than consonantal following segment. In other words, the interclass differences were greater when the (-t,d) cluster was followed by a non-consonantal segment (which included pause but, presumably, did not include liquids) than when followed by a consonantal segment. This was true for both monomorphemic and past tense words. Therefore, the effect of the following segment constraint as well as the overall rate of (-t,d) deletion distinguished the socioeconomic classes studied by Wolfram with the greatest difference shown in the effect of a non-consonantal segment.

Labov, Cohen, Robins, and Lewis (1968) studied (-t,d) deletion, among other variables, in African American speakers in Harlem. Their findings on the weighting of the constraints in the various age and socioeconomic groupings included in the study are discussed in Section 4.1.3 of this chapter. In general, however, they found that the higher the socioeconomic class, the more likely the speakers were to demonstrate an effect of the grammatical form constraint that was equal to that of the following segment constraint. (In the careful speech style, the middle class African American adults showed an effect of grammatical form which was stronger than following segment.) Like Wolfram, Labov et al. also found that the rate of (-t,d) deletion was inversely related to socioeconomic class. In the most favorable environment to deletion -- that of (-t,d) before consonant in monomorphemes -- the adolescent working and lower class peer groups deleted (-t,d) at well above the 90% level. The middle class adults simplified the clusters at the 79% level while the

Northern lower working class group deleted (-t,d) at a high rate in casual speech but at a lower rate in careful speech.

In summary, both studies support the notion that the overall rate of deletion increases as socioeconomic status decreases. They highlight different constraints as being most diagnostic of changes in social class, however, with Wolfram noting the phonological constraint to be the most indicative of social class differences and Labov et al. finding that it is the relationship between the constraints which is most influential.

4.2.2.3 The effect of gender on (-t,d) deletion

Gender differences are of interest to researchers in many fields, and sociolinguistics has a rich literature in this area, at least as it concerns adults. Labov (1990) discussed the dichotomy in findings in gender differences in variation. He pointed out that in situations of stable sociolinguistic stratification -- that is, when there is no language change in progress -- women tend to favor more standard speech forms. This is also the case in situations of change which occurs above the level of consciousness and is therefore subject to social pressure. On the other hand, in situations of language change coming from below the conscious level, women have been found to be leading the changes, using more extreme variants which are further from standard speech forms. Since (-t,d) deletion is a stable pattern of variation, one would expect women to use the more standard form, or to delete (-t,d) less frequently than men, and this is exactly what the research in this area shows. Both Wolfram (1969) and Neu (1980) studied the effect of gender on (-t,d) deletion.

Wolfram's subjects were white upper middle class speakers and African American speakers from the four socioeconomic classes listed above, aged 10 through adult. Neu looked at the speech of 15 white speakers, aged 19 to 53, from various geographical areas. Both found that males deleted (-t,d) at a higher rate than females.

4.2.2.4 The effect of age on (-t,d) deletion

The material on the acquisition of (-t,d) deletion has been discussed in Chapter 3. A review will be presented here as it applies specifically to the effect of age on (-t,d) deletion.

As noted above, Summerlin (1972) included African American and white lower class second graders as well as high school students and adults in her study of (-t,d) deletion in the rural South. She found that for the white children, there was no difference in deletion between the second graders and the high schoolers although both groups deleted (-t,d) at a higher rate than the adults (who were also middle class). The African American high school students, however, deleted (-t,d) more often than the second graders or the adults. Furthermore, although all of the children showed the grammatical constraint on (-t,d) deletion, neither the African American second graders nor the African American high schoolers demonstrated the phonological constraint. In general, Summerlin proposed that, as they grew older, the white children progressed toward the adult norm, whereas the African American children diverged from the white adult norm.

As previously described in Chapter 3 and reviewed earlier in this chapter, Guy and Boyd (1990) and Labov (1989a) also found indications of a developmental progression in the acquisition of the (-t,d) deletion rule, particularly as it pertains to semi-weak verbs such as *lost* and *slept*.

In general, the results of the studies on the effect of age on (-t,d) deletion are sparse and, again, underscore the need for a comprehensive study on the acquisition of the variable rule. Summerlin suggests that there is deletion going on in her second grade speakers and that the African American and white students may be heading in different directions as they acquire this rule. Both Labov and Guy and Boyd also find that young children delete final stops and are beginning the process of acquiring the constraints on this rule. In both cases, however, there was insufficient data to allow for individual analyses which are needed to further answer the question of acquisition of (-t,d) deletion.

4.4 Conclusions

(-t,d) deletion has been found in many studies over the years to be a stable, pervasive form of variation in English. Although the overall rate of deletion has been found to vary from social group to social group, the response of speakers to the various internal constraints is remarkably consistent. There are two exceptions to this consistency, both of which will be important to the current study. The first is the geographical dialect difference in the effect of following pause on (-t,d) deletion. The second is the formulation of semi-weak verbs as a separate morphological class. There is evidence to suggest that this rule is learned during the course of language acquisition. One of the charges of the current study will be to ascertain the presence and strength of these constraints in children for whom language is in the early stages of emergence.

CHAPTER 5: AN OVERVIEW OF THE VARIABLE (ing)

5.0 Introduction

This chapter will review the literature surrounding the variable (ing). Beginning with the social and stylistic variation which has been studied for the past thirty years, the chapter will then present the grammatical conditioning of (ing) and finally review the literature on its acquisition.

The primary contrasting variants of (ing) have been found to be [ing] and [ɪn]. Other variations, such as [iNk], exist, but these are infrequent and have generally been grouped with the [ɪn] variant. Woods (1979), however, noted the presence of a third variant [in] most prevalent in the middle class speakers he studied in Ottawa. He suggested that in most cases this variant is perceived by listeners as [iN]. Houston (1985) also reported the presence of [in] in British English and further suggested that the high tense front vowel /i/ influences the perception of the listener that [in] is equivalent to, if not indistinguishable from, [iN]. (ing) has been found to have a stable pattern of variation and one that is present in most dialects of English as will be seen in the studies of the variety of dialects reviewed below.

Labov (1966) discussed the history of the prestige attached to the two variants and noted that Krapp (1925) found that [ɪn] was the most common variant in New England as early as 1654. Dearborn's *Columbian Dictionary* (1795), however, lists [ɪn] as improper. Wyld (1936) stated that before the early 1800’s, what is now the [ɪn] variant of (ing) was universal. The change toward [iN] was caused by the influence of spelling, one that contributes to the stigmatization of the [ɪn] form which continues today.[5] Cofer (1972) documented the higher prestige value of [iŋ] by reading words containing both variants to 18 subjects. Of these, all 18 preferred *something* to *somethin'*, and 17 out of 18 preferred *working* to *workin'*. Only six of these informants admitted to using the [ɪn] form in either nouns or verbs, and none of them admitted using it in both nouns and verbs. Wald and Shopen (1985) interviewed men and women in Canberra, Australia and Los Angeles, California about the correctness of [iŋ] versus [ɪn]. They found that [iŋ] "emerged as unquestionably the standard" (p. 535) and that women were even more likely to prefer the [iŋ] form over the [ɪn] form than men.

5.1 Stylistic and social variation of (ing)

As stated in Chapter 4, stylistic and social factors in variation often overlap. It is often found that men, African American speakers, and lower socioeconomic groups are more likely to use less prestigious forms of variables, while women, white speakers, and higher socioeconomic groups use the more prestigious forms. These prestige variants are also more likely to be used in more formal interactions, such as testing, reading passages, or word lists. This was, in fact, the case with the adult literature on (-t,d) deletion reviewed in Chapter 4. The studies on (ing) production which will be discussed below are also very consistent with respect to stylistic and social factors and concur with findings on (-t,d) deletion.

---

[5]See Section 5.3 for an alternative variationist explanation for the change toward [iŋ] and the stigmatization of [ɪn] proposed by Houston (1985).

The quantitative study of the (ing) variable began with the work of Fischer (1958). He demonstrated both the stylistic and social variation of (ing) in 3- to 10-year-old children. His findings were that girls used more [iN] than boys. In addition, Fisher looked at the differences in (ing) production between "model" (e.g. school-oriented, especially well-behaved) boys and "typical" boys. He found that the typical boys were much more likely to use the [ɪn] form than the model boys. Finally, Fischer studied stylistic variation in one ten-year-old boy. He found that the boy used [iN] almost exclusively in a formal testing situation but used mostly [ɪn] in a less formal interview.

Labov (1966) looked at (ing) production in African American and white men and women and found that African Americans produced more of the [ɪn] form than whites. He studied the speech of his subjects in interviews which ranged over three styles -- casual, careful, and reading -- and found that as the formality of the language increased, so did the use of the [iN] form. He found no difference in (ing) production between those over and under the age of 20 and no gender differences.

Shuy, Wolfram, and Riley (1968) found race, gender, and social class differences in their study of (ing) production in adults. They noted that as socioeconomic status decreased, production of the [ɪn] form increased. In addition, African American speakers used more of the [ɪn] form than white speakers, and men used the [ɪn] form more often than women. They further noted that ethnic and gender differences were greatest above the social median. Below this level, only gender differences were statistically significant.

Cofer (1972) studied (ing) production in African American and white men, aged 22 to 44, in Philadelphia. He also found that African Americans used more [ɪn] than whites, and that working class speakers used more [ɪn] than middle class speakers. Stylistic differences were also consistent with other research on the topic in that speakers used more of the [iŋ] form in reading passages than in the less formal interview.

Trudgill's (1974) study of several variables including (ing) in Norwich English examined 60 male and female speakers ranging in age from 10 to 70. They were members of five socioeconomic classes from lower working class to middle middle class. He found that (ing) production distinguished all five of the social classes in that as social class decreased the percentage of [ɪn] use increased with the largest difference between the upper working class and the lower middle class. He also replicated earlier findings in terms of style: as the formality of the situation increased from casual speech to formal speech, reading passages, and word lists, the percentage of [iŋ] increased. Finally, he too found that men produced more [ɪn] than women.

Woods (1979) also found "strong stylistic variation" in a study of Ottawa English with speakers using less of the [ɪn] form in pictures, reading, and free speech than in minimal pairs and word list reading. He also found, however, than the variant [in] was most frequent in his study and that, with the exception of working class subjects, the speakers in general produced little of the [ɪn] form in free speech.

In 1985, Wald and Shopen summarized an earlier unpublished study of (ing) which took place in Canberra, Australia and Los Angeles, California. They focused on the effect of gender and addressee on (ing) production. In Canberra, men used more of the [ɪn] variant than women. The difference between the sexes was particularly notable when Wald and Shopen looked at the addressees of the speakers. They found that the men made a marked style shift toward the less formal [ɪn] variant when speaking to friends as opposed to non-friends. The women made no such style shift. All of the speakers, however, used more of the [ɪn] form when speaking to men than when speaking to women, but this shift was greatest for the women.

In Los Angeles, men used more [ɪn] than women with all addressees, but both sexes used more [ɪn] when talking to friends and family than to other addressees. Unlike in Canberra, in Los Angeles, women used more of the [ɪn] variant when talking to women than they did when talking to men although the findings for male speakers were consistent with those in Canberra. Findings in this study also included that the speakers over 40 years of age used more of the [iN] form than those under 40, who in this case were also their children. All speakers, but particularly female speakers, used more [iN] when talking across generations than when talking to their own generation. The younger generation also used more [iN] when talking to the opposite sex, but there was no such shift for the older generation. Finally, both groups used more [iN] when arguing about politics or morality topics than when carrying on a joking conversation.

Bradley and Bradley (1979) looked at (ing) production as well as other features in 30 adult speakers from Melbourne, Australia. They also found that men used more of the [ɪn] form than women. Unlike, Wald and Shopen, however, they did not find that addressee made significant differences in the rate of the (ing) variant produced.

In summary, the findings for social and stylistic variation in (ing) production are numerous and very consistent. Results of the various studies reviewed above uniformly show that African Americans and working class speakers used more of the [ɪn] variant than whites and middle class speakers. All of the studies but Labov (1966) found that males used more [ɪn] than females. Few of the studies looked at age differences, but of those that did, Labov found none while Wald and Shopen found that younger speakers used more [ɪn] than older speakers. Results of studies on the effect of style on (ing) show that the less formal the style of speaking, the more frequently the [ɪn] form is used. In general, the results show that the more standard variant [iɴ] is used more frequently by those who also are more likely not to delete (-t,d): women, white speakers, and those of higher socioeconomic groups. Also, as with (-t,d) deletion, [ɪn] is used more often in less formal interactional styles. (See Chapter 4 for a review of studies on (-t,d) deletion.)

5.2 Phonological conditioning and (ing) production

Studies of the (-t,d) variable show that speakers demonstrate the effects of following and preceding segments, and most, but not all of these, may be universally based. (See Chapter 4.) The phonological conditioning of (ing), on the other hand, appears to be restricted to regressive assimilation. Shuy, Wolfram, and Riley (1968), Cofer (1972), and Houston (1985) all noted that the presence of a following velar stop favored the [iN] variant, whereas the presence of a preceding velar stop favored the [In] variant. Similarly, the presence of a following alveolar stop favored the [In] form, while the presence of a preceding alveolar stop favored the [iN] form. For example, in *being good*, [iN] is favored, but in *making it*, [In] is favored. And, again, in *feeling tired*, [In] is favored, while in *beating up*, [iN] is favored.

5.3 The effect of grammatical form on (ing) production

Until 1979, it was assumed that the extralinguistic constraints were the primary ones operating on (ing). At that time, William Labov and his students in a research seminar at the University of Pennsylvania found during a survey of the Philadelphia speech community a grammatical effect on (ing). They noted that nominals and adjectives were more likely to be produced with an [iN] variant than were verbal categories. Gerunds, which share grammatical properties with both nouns and verbs were intermediary in promoting the use of [iN].

Houston (1985) further explored the grammatical effect on (ing) production in her study of (ing) in British English. She found an historical explanation for (ing) variation when she compared her modern data to that of c. 1450 reported by Moore, Meech, and Whitehall (1935). The 1935 study noted that in the fifteenth century, the verbal noun suffix *-ing* replaced the present participle suffix-*ind* in the South of England resulting in an identity of form between these two grammatical categories. Houston argues that this form identity was influenced by a perceived similarity in the suffixes caused by the high front vowel /i/. Her data also shows that today in this same area of southern England the probability of [iN] as opposed to [ɪn] was greater than .5. In the northern and peripheral areas of the country, however, the present participle -*and* was not replaced by -*ing* until much later, under the influence of written English. Houston found that the probability of [iN] in this area was less than .5. This geographical difference was most evident in the verbal categories, the same categories in which current studies find the most frequent use of [ɪn]. In this study as in the findings of Labov's research group, the more nominal words containing (ing) revealed more [iN] use than the verbal categories with the exception of monomorphemic nouns. These were usually pronounced with [ɪn] but were also generally restricted to *morning* and *evening*. In other words, Houston found a continuity between the present and past based on "categorical variation in the past being preserved as noncategorical variation in the present" (p. 287).

Houston's variationist explanation of (ing) production is in contrast to the previously held position argued by Wyld (1936) that the influence of spelling caused the change toward [iN] and the stigmatization of [ɪn]. She points out that Wyld's explanation does not account for the pronunciation of (ing) in Old English, nor does it discuss the grammatical effect on (ing) variation. She argues instead that the spelling change c. 1450 was representative of the change in pronunciation in the south of England.

The work of Houston on the historical significance of (ing) reveals a primary difference between the two variables included in this study. (-t,d) deletion is a grammatically and phonologically conditioned phonological process with a synchronic interpretation. That is, the loss of a final /t/ or /d/ in a weak past tense verb results in a loss of semantic content, and deletion rarely occurs in these words in adult speakers. There is no such loss of meaning when the process is applied to monomorphemic words, which more commonly undergo deletion. (See Chapter 8 for an alternative interpretation.)

In summary, although the number of studies showing a grammatical effect on (ing) production are fewer in number than those showing stylistic and social variation, the results of these studies are nonetheless consistent. More nominal grammatical categories utilized the [iN] form predominantly, while verbal categories favor the [ɪn] variant. Gerunds appear to be in the middle range between the two other categories. Houston’s work strongly supports an historical explanation for (ing) variation.

5.4 Acquisition of (ing)

As previously noted, Fischer (1958) found social variation in (ing) production in children aged 3 to 10. He did not break his findings down by age, however, so it is impossible to know whether the youngest children in his sample participated in this variation to the same extent as the older children. He also found stylistic variation in a 10-year-old boy but did not have sufficient data to address this issue with his other subjects.

Labov (1989a) looked at grammatical and stylistic constraints on (ing) variation in children. He found that the 7-year-old he studied had learned both the grammatical and stylistic patterns of (ing) variation, and a 6-year-old had acquired the stylistic but not the grammatical constraints. A look at limited data on a 4-year-old revealed no indications that he had mastered the constraints on (ing) production at all.

5.5 Summary

The above discussion demonstrates the consistency of the findings on (ing) production in terms of its stylistic, social, and grammatical effects. There are several studies from a variety of dialect areas documenting the existence of this variation and its social and stylistic effects. There are fewer studies on the grammatical effects on (ing) production, but it has been demonstrated in both American and British dialects. Fewest in number are the studies on the acquisition of this rule and its grammatical and stylistic constraints, and, at this point, it has not been demonstrated at all in preschool children. Therefore, the current study will look at larger numbers of children and at greater amounts of data to ascertain whether children are in the process of acquiring this grammatical rule and its constraints by the age of three or four.

CHAPTER 6: METHODOLOGY

6.0 Introduction

This chapter will explore in detail the methodology used in this study from the choice of the location for the fieldwork to the analysis of the data. Of particular importance is the modification of standard sociolinguistic interview practices for use with preschool children, which is described in Section 6.5.

6.1 The Philadelphia speech community

Labov (1989b) defines a speech community as "an aggregate of speakers who share a set of norms for the interpretation of language, as reflected in their treatment of linguistic variables." These commonalities are expressed by the fact that the speakers in a speech community share the same linguistic structural base. Although Philadelphia is a city of approximately 1,600,000 people, Labov points out that it is indeed a speech community in that speakers show consistent similarities with other Philadelphians and differences with outsiders in the distribution of words in phonemic categories, as in the case of the complex but widely used pattern of raising and tensing of short *a*, which is phonologically and lexically conditioned and appears to be unique to the Philadelphia area in its conditioning. Also relevant is the demonstration by Philadelphians of the effect of the dialect specific constraints on variable rules, such as that of following pause on (-t,d) deletion. [6]

---

[6]Labov (1980) notes that these generalizations about the Philadelphia speech community apply only to the white speakers of that area. African American and Puerto Rican speakers form their own speech communities and do not

Although the speech norms described above are common to the Philadelphia speech community as a whole, the city can also be accurately described as being made up of culturally distinct neighborhoods, often defined by the race, ethnic group, and/or socioeconomic status of the residents. The goal for this study was to choose an area of Philadelphia which would be likely to demonstrate strong application of the rules deleting (-t,d) and producing alveolar (ing) as well as similar reactions to the constraints governing their application.

Systematic linguistic change has been found to originate and become strongest and most advanced in the working and lower middle classes (Kroch, 1978; Labov, 1980). Neither of the variables in the present study involve sound change in progress. Nevertheless, it seemed advantageous to locate the research in a lower middle and/or working class neighborhood due to the fact that there would be fewer immigrants, more homogeneity, and less stylistic correction by the speakers who live there. South Philadelphia is an ideal site since it demonstrates exactly these socioeconomic characteristics. It is a primarily white neighborhood with many of its residents being of Italian background; and many families have been in this country for at least one generation. It is not uncommon to find grandparents, often born in Italy, parents, and children, if not in the same house, living within blocks of each other. Finally, South Philadelphia has been studied extensively over the years by the Language Change and Variation Project at the University of Pennsylvania under the

---

demonstrate the speech norms of the white community. Rather, African American Philadelphians share speech properties with blacks in other cities in the northern part of the United States.

direction of William Labov and by students in field methods courses at the same university, and its speakers have been shown to be among the most advanced in terms of use of the vernacular.

6.2 Entrance into the community

The decision to focus this work within the South Philadelphia area was the solution to only the first of several challenges. The second was finding children who were available to be tape recorded frequently over a period of a few months in order to obtain sufficient data for analysis. To begin the search process, I contacted several South Philadelphia day care centers listed in the Yellow Pages of the Philadelphia telephone directory. I requested information similar to that which might be requested by the parent of a prospective student. Of the greatest interest was the size of the center, so a sufficiently large subject pool could be obtained, the ages of the children, and whether or not most of the children resided in the neighborhood. The latter question was asked in an effort to keep socioeconomic factors as equal as possible across subjects. It was also important for the children to be members of the same speech community, even though the rules themselves are used widely among speakers of English. The following segment constraint on (-t,d) deletion is geographically dependent, however, and I wanted to be able to assess whether or not the children had acquired it.[7]

---

[7]As was noted in Chapter 4, Guy (1980) found that a pause following a (-t,d) consonant cluster encouraged final stop deletion in New Yorkers and in African American speakers from more than one city but inhibited it in white Philadelphians.

I found that all of the day care centers I contacted had enrolled a large enough number of 3- and 4-year-old children for the present study. Those centers deep in South Philadelphia were most likely to have children from their immediate neighborhoods attending them. The centers nearer to Center City, Philadelphia's business district and South Philadelphia's neighbor to the north, were eliminated since they seemed to have a greater diversity of students including those whose parents worked in Center City but lived elsewhere in the city, the surrounding suburbs, or New Jersey.

The director of the first remaining center I called, Kids' Land, agreed to allow me to tape record the children in her care. Kids' Land is a privately owned day care center run by the director, Donna,[8] and her husband. It is located on the first floor of two row houses in which the adjoining wall had been opened up into an archway to connect the two rooms. There are no other walls or dividers, except those separating the two bathrooms, so the effect is one of two large, open but connected spaces. I first visited Kids' Land in December, 1988, and found the center filled with Christmas decorations, including an almost-life-sized creche, all made by Donna. It was clear upon observing the center and talking with Donna that she carried her role as owner/director far beyond a supervisory or administrative one. In addition to decorating the center, she created most of the curriculum including writing songs for various holidays and special events. She often worked with the children directly, and her desk was located just inside the front door of Kids' Land, so she

---

[8]All subjects, their teachers, relatives, and friends, as well as the day care center itself, have been given pseudonyms to protect their privacy.

was always present and accessible to the staff and children. She appeared to know all of the children and most of their parents well. When I talked to the parents later, they often stated that they first sent their children to Kids' Land because they were friends of (or, in some cases, related to) Donna and her husband.

In addition, there were usually three or four teachers present. Staff turnover was frequent, and it was difficult for me to keep track of teachers while I was there. Those I met, however, were also from the South Philadelphia area. Also present from time to time at Kids' Land were Donna's husband, whose presence was noted particularly when something needed to be fixed, and their two school-aged children who spent the day at Kids' Land when their school was not in session. They helped with the younger children, and Donna's daughter was a part of my taping sessions on occasion. Enrollment at the center varied from day to day but was generally between 18 and 23 children. Children could begin attending Kids' Land when they were toilet trained, and they usually continued there until they "graduated" into kindergarten.

I tape recorded the children at Kids' Land from mid-January through April, 1989. I usually arrived between 8:30 and 9:00 a.m. and set up my equipment in a corner of the room. I recorded the children throughout the morning until nap time, which took place just after lunch.

The children's morning began with free play and snack. This was followed by circle time, for which the children were divided into two groups by age. During circle, they sang songs, played games, listened to stories, and learned to recognize letters, numbers, and their own written names. Circle always ended with the recitation of the Pledge of Allegiance, for which the flag was held by one of the children. Lunchtime was also divided into two sessions with the group not eating lunch spending its time watching children's videotapes on television.

Since there was no wall between me and the rest of the activity, I was able to watch the goings on of the center while I was recording my subjects. Conversely, they were also able to watch me -- a practice I in no way discouraged. I even noticed, on occasion, one of the games I played with the children turning up in a slightly altered form in circle time. This process seemed to make my presence comfortable for both teachers and parents. On the other hand, background noise occasionally became a problem, and taping would be suspended briefly until a particularly loud activity was over.

Toward the end of my time at Kids' Land, I began contacting the parents, some of whom I had met at the center, for background information. Eight of these I interviewed in person in their homes, and the remaining nine I attempted to contact by telephone.[9] Of these, only two refused to be contacted. One mother told Donna she had just had an unlisted telephone number installed and did not want the number released. In the other case, a child's mother gave her permission to Donna for me to call her. When I called, however, the child's father answered the telephone, screened the call, and was

---

[9]Donna suggested the procedure of her asking the parents for permission to release their telephone numbers to me. I would then call them, since it would have been inconvenient for them and difficult for me to try to catch them at the center as they dropped off their children. This was the method we used.

hostile and suspicious. Since I was unable to allay his fears, I did not attempt to make contact again.

From those I successfully reached, I obtained a history of their places of residence as well as a brief description of their child's language acquisition. I also asked about their feelings on teaching their children language and about any other persons who were important influences on their children.

6.3 The subjects

Seventeen children, 11 girls and 6 boys, were tape recorded. At the beginning of taping, they ranged in age from 3 years 2 months to 4 years 11 months. The mean age of the children was 3 years 11 months. None of the children appeared to have speech or language learning difficulties with the possible exception of Callie who had difficulty producing consonant clusters and an unreliable past tense. Therefore, her data were eliminated from the analysis of (-t,d) deletion. All of the children but one were living at the time with at least one of their parents. The one exception (Jenny) lived with her paternal grandmother and saw her father when he visited their home. Of the 15 children for whom I was able to obtain parent information, all were raised in South Philadelphia. Thirteen of the children had parents who were raised in the greater Philadelphia area; ten of them were born and raised in South Philadelphia itself. Of the two remaining children, Gia's mother was raised in a family in diplomatic service which moved frequently, while her father is a life-long Philadelphian. Mike's parents were both born and raised in Italy and moved to Philadelphia as adults. All of the children but one were white. Kent was the child of a African American father and a white mother. He was raised solely by his mother in a white neighborhood, however, and identified himself as white. The families were for the most part working to lower middle class, and the occupations of the parents included construction worker, homemaker, word processor, optician, secretary, carpenter, restauranteer, factory worker, retail manager, and tailor.

The above description of the subject group has centered on the families of the children. This is different than many studies of variation, particularly those involving young people, which focus on peer group (e.g. Labov, 1966). For children of the preschool age, however, the family, including the extended family, is more important and more stable than peer group. This is particularly true for this group of children, many of whom entered the day care center within one year of their interview. The description of their personal history, therefore, supersedes that of their social peer network.

6.4 The variables

There are two variables examined in this study: (-t,d) deletion and (ing) production. The history of research on these variables is discussed in Chapters 4 and 5. This section will define these variables, the rules governing their use, and the constraints that affect these rules.

6.4.1(-t,d) deletion

The (-t,d) variable is defined as a word-final consonant cluster ending in /t/ or /d/. Unreleased stops and flaps were counted as realized stops. (-t,d) clusters followed by a dental segment or alveolar stop were omitted from the data since the presence of these homorganic following segments make it impossible to determine whether or not the (-t,d) was produced. Also omitted were (-t,d) clusters preceded by /n/ and followed by a vowel since this combination results in the neutralization of (-t,d). Finally, the word *and* was eliminated from the data because (-t,d) was found to be deleted from this word at a rate approaching 100% regardless of the phonological environment.

As discussed in Chapter 4, (-t,d) is variably deleted from word-final consonant clusters, and this deletion is favored or disfavored by the presence or absence of various factor groups: following segment, grammatical form, syllable stress, etc. Most of these factor groups will be examined in the current study, so it was necessary to code the data as to the presence or absence of (-t,d) in each case. The factor groups under consideration in this study included the following:

Linguistic factors:

1. Following segment: The tokens were coded with regard to whether they ended with an obstruent, glide, liquid, vowel, or pause.
2. Grammatical form: The categories for grammatical form included contractions ending in *n't*, monomorphemic words (e.g. *nest, left foot)*, weak past tense verbs (e.g. *talked, hopped*), semi-weak or ambiguous verbs (e.g. *slept, told*), and participles (e.g. *the baked cake*).

3. Syllable stress: The data were coded as to whether the (-t,d) cluster appeared in a stressed or an unstressed syllable.

4. Presence of a preceding consonant: Words such as *first* which end in a three-consonant cluster were coded as having a preceding consonant.

Extra-linguistic factors:

1. Addressee: The tokens were coded as to whether they were included in utterances the child addressed to an adult, a child, or an inanimate object such as a toy or puppet.

2. Style: As discussed in Chapter 4, (-t,d) deletion has often been found to be affected by the style of the interaction. Specifically, the more informal styles (e.g. general conversation, especially the telling of narratives) have been associated with a greater probability of (-t,d) deletion than more formal styles (e.g. reading, word lists.) Coding for style in this data was somewhat problematic since the classification of language styles often used with adults (e.g. narrative, conversation, reading, word list) was not applicable for the children. They were unable to read and did not, as a rule, tell narratives. The attempts to elicit stories from them in the story telling or pretend "book reading" activities did result in much useful data, but not in the informal style associated with adult narrative speech. Rather, the most commonly heard style

during these book activities was the "reading intonation" style described by Scollon and Scollon (1981) in their study of pre-literacy in young children. Nevertheless, an attempt was made to see if any style shifting which might occur with a change in activity would have an effect on (-t,d) deletion or (ing) production. Therefore, the data were coded as to whether they took place during a "book reading" activity, a picture-naming game, role playing (e.g. talking for a puppet or other character), which took place during an imaginative play activity, or other.

Sutton-Smith (1971) noted that at approximately three years of age the child expands upon her previous dyadic play interactions between herself and a doll, for example, and includes other characters so that "a plurality of relationships may be represented" (p.301). The child may play several roles and add imaginary characters. In the present study the use of a Sesame Street toy with manipulable characters encouraged the children to pursue this imaginative play and allowed them to talk to and role play several characters. The utterances were then divided by activity as a way of examining style to see if the role playing resulted in a more informal style than the presumably more formal picture naming or book reading activities.

3. <u>Subject:</u> Each subject's tokens were coded separately so individual differences could be examined.

4. Sex of subject: All tokens were coded as to the sex of the speaker.

5. Age of speaker: The speakers were divided into four groups: 3-year-olds, 3 1/2-year-olds, 4-year-olds, and adults.

6. Mean length of utterance (MLU): The child subjects were divided into two groups depending on the mean length of their utterances by morpheme. MLU was first discussed by Brown (1973) who emphasized that children's rates of language acquisition vary widely. He further stated that "two children matched for MLU are much more likely to have speech that is, on internal grounds, at the same level of constructional complexity than are two children of the same chronological age" (page 55). Following his guidelines, the morphemes in 100 utterances were counted for each child.[10] The total was then divided by 100. The children's MLU's ranged from 3.43 to 6.31 morphemes per utterance, and the group was divided into two parts between the MLU's of 5.33 and 5.38, resulting in a Low MLU group of nine children and a High MLU group of eight children as shown in Table 6.1.

---

[10]See Appendix A for Brown's rules for the calculation of mean length of utterance.

Table 6.1: List of child subjects with their mean lengths utterance in morphemes (MLU) and their ages.

Name	*MLU*	*Age*
LOW MLU GROUP		
Cindy	3.43	3-3
Diane	4.12	4-6
Mike	4.52	3-4
Marie	4.65	3-2
Micky	4.78	3-5
Evan	5.01	3-10
Rhea	5.17	3-11
Callie	5.30	3-11
Erin	5.33	3-3
HIGH MLU GROUP		
Jeanie	5.38	3-9
Zak	5.53	4-11
Mira	5.59	4-3
Jenny	5.82	3-11
Kent	5.82	4-6
Danny	5.90	4-10
Shelly	6.29	4-9
Gia	6.31	3-11

There was no significant correlation between age and MLU (Kendall coefficient of concordance=.794, df=16) or between sex and MLU ($mean_{girls}$=5.22, $mean_{boys}$=5.26, t=.11, df=l5) in this sample. The relation between the age and MLU has been discussed often in child language acquisition literature with varying results. Brown (1973) noted that MLU is a better predictor of grammatical development than age. Miller and Chapman (1981) found, however, that age and MLU were correlated for 123 middle and upper middle class children between the ages of 17 months and five years. This correlation was particularly strong up to the age of four years.

On the other hand, Klee and Fitzgerald (1985) studied eighteen 2- to 3-year-old children and found that MLU did not correlate significantly with age. They also found that MLU did not significantly discriminate children's profiles of syntactic diversity at the clause or phrase level. It did, however, correlate significantly with increased use of bound morphemes. The authors noted that their study focused on children in a more restricted age range than that of Miller and Chapman, and this factor, resulting in a more homogeneous group and decreased variance, may have accounted for the differences in results. This hypothesis is supported by Miller and Chapman's reanalysis of the deVilliers and deVilliers (1973) data which showed a significant correlation between age and MLU when their full subject pool (ages 16 to 40 months) was included but a chance rating when age was restricted to 21 to 33 months.

Given these findings, it is not surprising that there were no significant findings between age and MLU in the current study of 3- and 4-year-olds. This is particularly true since even Miller and Chapman found that variability in predicted MLU increased with age, and the relation between age and MLU decreased at age 4 and above. In addition, this negative finding between MLU and age and that of Klee and Fitzgerald between syntactic diversity and MLU underscore the point that mean length of utterance more accurately represents a differential acquisition of forms than a linear overall increase in complexity.

6.4.2(ing) production

The variation of (ing) has been shown by Houston (1985) to be an historical alternation reflecting a partially complete merger between two morphemes which were originally distinct in English. (See Chapter 5 for a more detailed discussion of this partial merger and its consequences in modern English.) This rule-governed alternation is describe by Houston as follows: "A [syllable final] back consonant (nasal) variable goes to minus back in the environment of a preceding unstressed vowel" (page 24). In the present study (ing) varies between the forms [ɪn] and [iN]. (There were no instances of [in] among the tokens in this study.) As in all studies of this variable in Philadelphia, *everything* and *anything* were omitted because they are produced categorically as ending in [iN]. The data were coded for one linguistic factor group and six extra-linguistic factor groups as follows:

Linguistic factor:

Grammatical status: The tokens were coded as to whether they were progressive verb forms (e.g. *running*), complements (e.g. *He's finished eating.*), verbal adjectives[11] (e.g. *swimming pool*), and nominals (e.g. *wedding, morning*). *something* and *nothing* were coded as a separate factor.

---

[11]The unique characteristics of the preschool-age child data necessitated some modifications in the more common coding techniques for (ing). There were no instances of true adjectives, such as *exciting* or *frightening*. Therefore, verbal adjectives made up the closest thing possible to an "adjective" class, and were coded separately from gerunds functioning as nominals, as in "Swimming is fun!"

Extra-linguistic factors:

The extra-linguistic factors examined for the (ing) variable were the same ones described above for (-t,d) deletion.

6.5 The play-interview method

As discussed in Chapter 1 and elsewhere, one of the difficulties encountered by Guy and Boyd (1990) and Labov (1989a) was that there was too little data available from the youngest children in their studies to allow individual variable rule analyses. Guy (1980), in discussing this issue, suggests a goal of 30 tokens per factor in the smallest factor. Since the factor groups for (-t,d) deletion and (ing) production have up to five factors each, it would seem reasonable to work toward the goal of approximately 150 tokens for each variable. While it may be reasonable to expect this many tokens in an adult interview of approximately two hours in length, it soon became apparent in interviewing children that it would take much more time to achieve a number of tokens even close to the goal. The challenge was to gather the maximum amount of data in the fewest possible taping sessions. Of course, this is always one of the goals of sociolinguistic interviewing, and Labov (1984) describes in detail techniques which work quite well when used with adult subjects. These include questions on a variety of topics including common sense solutions to problems, danger of death, and the speaker's neighborhood. The techniques are designed to elicit large amounts of speech as well as narratives in an informal style so as to gather instances of the day-to-day vernacular spoken by the subjects. The questions are usually supplemented by more formal techniques such as reading passages, word lists, and minimal pair lists to obtain individual specific tokens needed in the study of a particular variable. Needless to say, much modification was necessary to adapt this method to children.

Researchers in the field of child language acquisition have been collecting spontaneous speech samples from children in cross-sectional and longitudinal studies for years, and it seemed appropriate to look to this body of research for help in defining the methodology for the current project. The technique often used in child language studies is the play session with variation in the duration and frequency of the sessions. The play includes age appropriate toys and can involve the researcher, a parent or other familiar adult, or, sometimes, another child as play partner. In general, the play sessions are kept as unstructured as possible in order to elicit spontaneous and typical language from the child. To take two classic examples, Brown (1973) tape-recorded play sessions with three children for a minimum of two hours per month over a period of from one (Sarah) to five (Adam and Eve) years. The basic schedule for Adam and Eve, who were 18 months old at the start of the study, involved one two-hour session per month. Sarah was seen weekly for half-hour sessions from the age of 27 months. deVilliers and deVilliers (1973) undertook a cross-sectional study of 21 children between the ages of 16 and 40 months. They saw each child for two, 1 1/2 hour play sessions.

While certain aspects of the play session were found to be extremely useful in the present study, there were differences related to the purposes of the respective studies that required adaptation of the play session methodology. First of all, the amount of data required and obtained by the studies described above was less than that needed for individual variable rule analysis. Both of the studies examined mean length of utterance and morpheme acquisition. Mean length of utterance was calculated using 100 utterances -- usually very easily obtained in a singe, half-hour session. The individual morphemes of each speaker were included in the analysis if there were at least five obligatory contexts for a given morpheme (e.g. "two coat" or "I want some cookies" for regular plurals). In contrast, 150 (-t,d) or (ing) tokens per child were the goals of the present study. Consequently, more sessions were needed for this cross-sectional study than were required by that of deVilliers and deVilliers or at each age for the children studied by Brown.

The second difference between studies of categorical and variable rules is that, for the latter, large concentrations of tokens are required for specific grammatical forms (e.g. regular and semi-weak past tense verbs) and, sometimes, examples of particular lexical items (e.g. *cooking* as a noun, verbal adjective and progressive verb). Therefore, the sessions used in the present study needed to be somewhat more structured to allow for the introduction of activities or toys which would elicit the production of the desired tokens.

The solution to this methodological dilemma was a group of sessions which combined the play aspects of the language acquisition studies with the increase in the structure of activities of the sociolinguistic interview. In addition, increased numbers of interviews were needed to provide the greater amount of data, which is more difficult to obtain from children than from adults and is required for variable rule analysis. The result was the play-interview session which is described below.

The children were interviewed, one or two at a time, for 20 or 45 minutes each, depending on the length of the tape and the child's attention span. The children and I were seated around a child-sized table with the toys or other props placed in front of us. A Nagra reel-to-reel tape recorder was used with a Sony ECM 50 lavaliere microphone worn by the child. When two children were recorded at the same time, one of them wore the microphone. This equipment produced very good sound, and it was found to be important to have sound equipment which was as high quality as possible to capture the small details of the children's articulation in spite of the noise in the background.

The play-interview sessions began with conversation: first, with my asking questions to gather background information; then, as we got to know each other better, with the children volunteering accounts of their activities at home or at school. It was clear, however, that additional techniques would be necessary to gather the large amounts of speech needed for this study.

Many activities, such as shopping with play food and money, playing with a doctor kit, and putting puzzles together, were tried in an attempt to elicit large quantities of informal speech. The most successful activities, however, were those which combined play-acting with manipulable materials. For example, a very useful activity was that of playing with a Sesame Street house which came with furniture, props, and characters. The children talked while setting up the toy, and then described the action and talked to and for the characters as they played. Specific lexical items, such as those with final (-t,d) clusters could be encouraged by having those objects in the house (e.g. *Big Bird's nest*). Also successful in eliciting large amounts of interaction was a set of identical toy telephones.[12] This activity was particularly good when used with two children, although it was also used with one child who could "talk on the telephone" to me.

Slightly more structured activities were tried as well with good results. Most useful were two "reading" activities. The first involved giving a child a picture book of a familiar fairy tale such as *Goldilocks and the Three Bears* or *Little Red Riding Hood*. The children were asked to "read" or tell the story. The pictures helped the children add detail to the story which might have been forgotten without the pictures, and the presence of words on the pages did not interfere with their creative rendering of the stories, since none of the children could actually read. The second reading activity was a variation of the picture book except that the children were handed a completely blank book and asked to pretend to read a story. Children who were at first hesitant to begin were encouraged by the suggestion that the story could be one that they knew or one about themselves. Both story activities were particularly effective at getting instances of past tense, necessary for the (-t,d) deletion analysis, and the present progressive verb form, necessary to the

---

[12]It was found to be important that the telephones be identical to eliminate long stretches of arguing about who got which telephone. Argument is, of course, usable interaction but tended to be very repetitive.

(ing) analysis. These forms were further encouraged by the prompts, "Then what happened?" or "What's she doing now?"

The most formal portion of the play-interview involved the naming of picture cards to elicit productions of words I was unable to get in any other less formal way.[13] This activity was made more entertaining for the children by hiding the pictures in a cloth bag. The children retrieved a picture and named it. When all of the pictures were taken from the bag, the children attempted to toss the cards into the bag after naming the picture again. This last had the additional benefit of eliciting instances of "*I missed!*" for the (-t,d) deletion analysis.

Finally, puppets were used often throughout the play-interview sessions. The reason for this was two-fold. First, they seemed to make the children more playful and, it was hoped, would elicit more informal interaction from them. Second, the names of the puppets were carefully selected to contain target sounds. In this case, the desired phoneme was short *a*, and the puppets names included *Sally*, *Janet*, and *Allan*. (See footnote #10.)

The above techniques and four months of taping resulted in 138.5 hours of tape for the child speakers. Enough data was collected on most of the children to allow for the individual analyses. The amount of time each child was tape recorded ranged from 2.5 to 7.5 hours. Those with the shorter amount of time were often unavailable for interviewing due to absence from day care. In

---

[13]This last activity was very seldom necessary to gather data for the present study, but it was found to be very useful in collecting instances of the Philadelphia short *a* pattern for another project. (See Roberts and Labov, 1992.) It is included here to allow for replication and/or elaboration of the methodology.

almost all cases, approximately 4.5 hours of tape or 9 to 14 interviews were necessary to obtain the amount of data often elicited in a single adult sociolinguistic interview. Even with the above techniques and the number of tapes recorded, however, there were some children for whom individual analyses were not possible. There was also one factor in particular for which more data was needed. The semi-weak or ambiguous verbs proved to be the most difficult to elicit, but, nevertheless, there were enough collected to make some generalizations possible regarding this verb form.

As previously mentioned, parents of eight of the children were interviewed. When both parents were living in the home, they were both invited to participate in the interview. In all but two cases, however, the interview was done with the mother alone.[14] The purpose of these interviews was two-fold. One reason was to collect background information, and for this reason, questions were focused on the child's development and the parents' attitudes toward parenting. The second reason was to provide comparison data for the findings from the variable rule analysis of the children's speech. Because of the focus of the questions, the range of styles was usually less varied and the interviews often shorter than the more standard sociolinguistic interviews. Consequently, the amount of data was also often less. However, there was still enough data collected in most cases to provide a basis of comparison, particularly when findings from other adult studies of (-t,d) deletion and (ing) production were

---

[14]In one of the exceptions the mother and father participated in the entire interview. Unfortunately, however, the father rarely talked. In the other, the father came down for a ten minute interview during a break in a hockey game. I found that one of my greater methodological errors was to schedule my parent interviews during hockey play-off season.

also considered. Table 6.2 constitutes a review of all of the speakers and their interviews, showing the names of the subjects, the total number minutes a speaker was tape recorded, and the number of (-t,d) and (ing) tokens for each as well as the ages and MLU's of the children.

Table 6.2 Amount of (-t,d) and (ing) data gathered for each child.

Name	Age	MLU	# Min.	# Tokens (-t,d)	# Tokens (ing)
Cindy	3-3	3.43	400	160	63
Diane	4-6	4.12	280	77	46
Mike	3-4	4.52	280	72	86
Marie	3-2	4.76	220	96	74
Micky	3-5	4.78	280	109	91
Evan	3-10	5.01	380	102	73
Rhea	3-11	5.17	240	43	59
Callie	3-11	5.30	460	n/a	75
Erin	3-3	5.33	180	44	48
Jeanie	3-9	5.38	400	173	174
Zak	4-11	5.53	300	181	151
Mira	4-3	4.30	200	93	112
Jenny	3-11	5.82	280	169	109
Kent	4-6	5.82	160	63	41
Danny	4-10	5.90	300	154	106
Shelly	4-9	6.29	360	169	158
Gia	3-11	6.31	340	127	115

Parent's Name	Child's Name	#Min.	#Tokens (-t,d)	#Tokens (ing)
Dee	Gia	25	62	39
Kay Ann	Jenny	40	77	44
Marianna	Callie	40	35	39
Carla	Evan	80	114	79
Donna	Danny	80	94	125
Lois	Cindy	40	29	38
Mary	Jeanie	40	26	23

Debra	Shelly	80	140	92
Jack	Gia	15	17	10
Tom	Jeanie	40	9	8

6.6 Variable rule analysis

After the data were coded in the categories described in Sections 6.4.1 and 6.4.2, they were analyzed using Goldvarb 2.0 (Rand and Sankoff, 1990). This multivariate analysis program which uses the method of maximum likelihood emerged from an older version using an additive algorithm. The program can be used with linguistic and extra-linguistic factors and results in probabilities of rule acquisition in the presence of a given factor or factors. Prior to this time, percentages were used to demonstrate variation in language, but this method failed to account for the fact that more than one factor can simultaneously affect the probability of rule application on a given variable. Standard analysis of variance methods could not be used with naturally occurring linguistic data due to the unevenness of the distribution of the tokens among the factors. Since 1969, developments have taken place in the methodology including a movement from the additive algorithm to a multiplicative one introduced by Cedergren and Sankoff (1974). This method of variation analysis was called Varbrul, a name which covers the analysis of variation by any of several multivariate analysis programs based on the maximum likelihood algorithm. The Goldvarb program is a version of Varbrul formulated for use with Macintosh computers.

6.7 Summary

This chapter has examined the methodology used in the current study of (-t,d) deletion and (ing) production in preschool children. The fieldwork for the study took place in a day care center in a working to lower middle class area in South Philadelphia where seventeen 3- and 4-year-old children were recorded over a period of about four months. Extreme modification of the standard sociolinguistic interview practices was necessary to adapt them for use with young children, and the resultant procedure might best be termed a play-interview session. It consists of activities of varying formality designed to elicit the maximum amount of speech from the children in the minimum number of taping sessions. Eight of the children's parents were also interviewed as a means of obtaining background information as well as providing a comparison for the child data. The data were coded and analyzed using the Goldvarb 2.0 program of multivariate analysis for linguistic data (Rand and Sankoff, 1990.)

CHAPTER 7: RESULTS AND DISCUSSION OF THE ANALYSIS OF (-t,d) DELETION

7.0 Introduction

The purpose of this chapter is to outline the results of the analysis of acquisition of the (-t,d) deletion rule by sixteen 3- and 4-year-old children and to discuss these results as they relate to both sociolinguistic and psycholinguistic research. In addition, three specific questions will be examined in the discussion. The first is whether the probability of (-t,d) deletion and the effect of constraints on it can be explained as a natural performance constraint on the articulation of the consonant cluster. In other words, can ease of articulation of the final /t/ or /d/ in a cluster in a given environment account for its presence or absence? Work on (-t,d) deletion in adults has already addressed this question. Labov (1967) suggested that the ordering of the following segment constraint of consonant > liquid > glide > vowel[15] may be universal and functionally based whereas the effect of following pause is based on dialect. Guy (1980) provided confirmation for this suggestion when he showed that following pause inhibited deletion for Philadelphians but promoted it for New Yorkers. This question has never been addressed, however, using the speech of very young children. It remains to be seen whether children learn this type of dialect specific feature while they are learning the language itself. If so, the results will speak strongly against the hypothesis that

---

[15]The symbol ">" is to be read as "favors deletion more than."

variation can be accounted for by natural and universal performance constraints on language.

A related question is whether these results can be accounted for by a universal tendency to preserve meaning in speech. Several past findings argue against this possibility. First is the fact that participles, which have redundant past tense marking, show almost identical deletion levels to weak past tense verbs, which are not redundantly marked. Second, van Hout (1993) reported that in Dutch /t/ is more likely to be deleted in second or third person present tense verbs, in which it is the only case marker, than in past participles, which contain redundant marking, and monomorphemes. Finally, the results of Guy and Boyd (1990) contrasting deletion in weak past tense and semi-weak verbs suggest that preservation of meaning cannot account for findings on (-t,d) deletion since the different analyses of their subjects of different ages imply that they cannot be participating in such a universal tendency in regard to these grammatical forms. The present study will further explore this issue with pre-school speakers.

The second question is, if the children demonstrate similar patterns of (-t,d) deletion to their parents and other adults, are they actually learning variable rules and their constraints, or are they simply matching the probabilities of individual words they hear from the adults around them? Research using adult speakers as subjects cannot address this issue, but the work of Guy and Boyd (1990) and Labov (1989a) have both commented on it. If, as their work indicates, the younger subjects differ from the adults in their response to the grammatical form constraint and its effect on deletion in semi-weak verb, they cannot be merely matching the probabilities used by their parents. Labov's subject was 7 years old, however, and Guy and Boyd suggest that the (-t,d) segments in the semi-weak verb class may not be there at all for their youngest subjects. In addition, as previously discussed, the amount of data in both of these studies, particularly as regards this grammatical form, was limited. It is important, therefore, to examine closely the (-t,d) deletion of very young children to see if they do analyze these verb differently than the adults from whom they are learning language.

7.1 Effect of following segment on (-t,d) deletion

Guy (1980) found that the adults in his study demonstrated the following effect of following segment on deletion:

consonant > *liquid* > *glide* > *vowel*

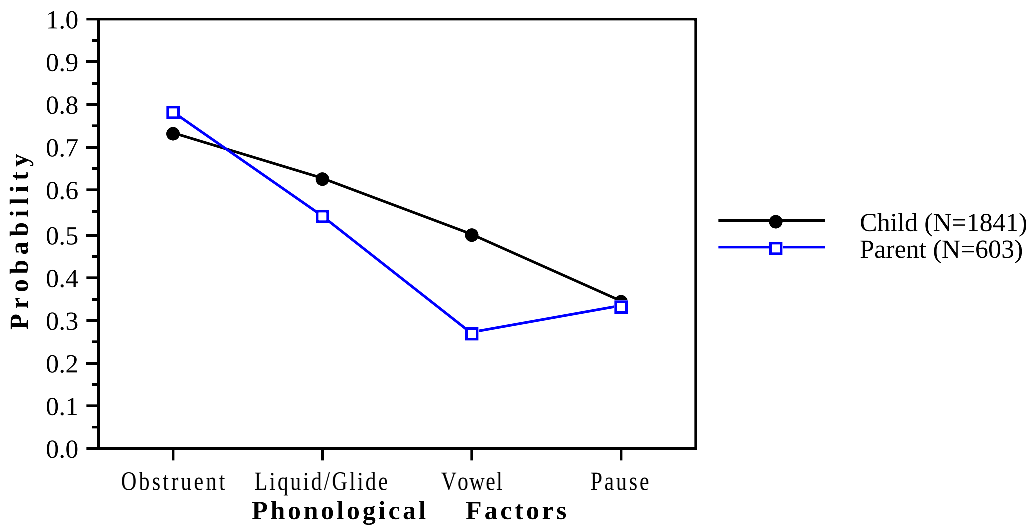
These findings were true of adults regardless of geographical area. As discussed in Chapter 3 and the previous section of this chapter, the effect of pause varies by geographical area. For Philadelphia speakers following pause exerted an inhibitory effect on deletion making the continuum as follows:

consonant > *liquid* > *glide* > *vowel* > *pause*

As can be seen in Figure 7.1, the adults in the present study followed a similar pattern although not all differences were significant. Specifically, there was no significant difference in the effects of following liquid, glide, and vowel. Differences between following obstruent and the combined factor group of following liquid, glide, and vowel were significant at the level $p<.01$. The same is true for the difference between the liquid-glide-vowel group and following pause.

As is also shown in Figure 7.1, the children in this study were significantly more likely to delete /t/ and /d/ following a consonant than a vowel. Liquids and glides in this case were combined since for the children, like the adults, no significant difference between them was noted. These children also conformed to the Philadelphia pattern of disfavored (-t,d) deletion preceding a pause. The difference between following obstruent and liquid/glide and between liquid/glide and vowel were significant at the level $p<.05$. The difference between following vowel and following pause was significant at the level $p<.01$.

Figure 7.1: Effect of following segment on (-t,d) deletion.



The group results show that the children have acquired the following segment constraint on (-t,d) deletion, including the geographically determined following pause constraint. In studying language acquisition and sociolinguistics, however, it is important to look at individual behavior as well as that of the group. Therefore, the next figures and discussion will present the deletion patterns of each child and explore the extent to which they are similar to the patterns of the group as a whole. The results for individual children in the case of following segment were striking in their conformity with the adult results. First, since there was not enough data per factor group for each child to run individual variable rule analyses for all of the children, the individual data are first displayed by percentages in Table 7.1. For the obstruent factor, the range of percentages of deletion was from 37% to 79%, whereas for vowels it was 0% to 47%, and for pause, 7% to 40%. The area of overlap between the range of percentages for obstruents and that of the range of percentages for following vowel and pause is very small. Only four of the children had percentages in that area of overlap. Zak and Evan had percentages of deletion following an obstruent of 37% and 43% respectively, and Cindy and Jeanie had 38% and 47% deletion preceding a vowel. Without exception the children demonstrated the same pattern as the adults of deletion being more favored before an obstruent than a vowel or pause.

Table 7.1: Effect of following segment on child subjects: and glide categories combined.

Name	Age	MLU	Obs. N	Obs. %	L/G N	L/G %	Vowel N	Vowel %	Pause N	Pause %	Total N
Cindy	3-3	3.43	25	68	16	69	47	38	72	21	160
Diane	4-6	4.12	18	56	14	57	10	0	35	40	77
Mike	3-4	4.52	4	75	4	50	10	30	61	13	79
Marie	3-2	4.76	19	63	13	54	14	14	50	14	96
Micky	3-5	4.78	22	64	12	25	26	8	49	12	109
Evan	3-10	5.01	21	43	18	28	11	0	52	31	102
Rhea	3-11	5.17	12	67	8	50	8	13	15	27	43
Erin	3-3	5.33	11	55	11	0	5	20	21	38	48
Jeanie	3-9	5.38	48	79	24	50	36	47	75	23	183
Zak	4-11	5.53	38	37	36	22	36	19	71	27	181
Mira	4-3	5.59	27	52	18	50	8	25	40	40	93
Jenny	3-11	5.82	35	54	16	69	45	29	73	21	169
Kent	4-6	5.82	21	67	18	28	10	30	14	7	63
Danny	4-10	5.90	37	62	21	52	32	25	64	28	154
Shelly	4-9	6.29	31	55	21	52	30	20	87	33	169
Gia	3-11	6.31	30	50	28	57	15	0	54	30	127
Mean %			59.19		44.56		19.86		25.31		115.81
Standard Deviation			11.11		18.70		13.70		10.18		48.07[16]

These results are shown graphically using probabilities in Figure 7.2. In this case the findings for only thirteen children are shown since for the others, the total amounts of data were too small for individual analyses. In variable rule analysis, interpretation of results relies not on the specific value of the probabilities

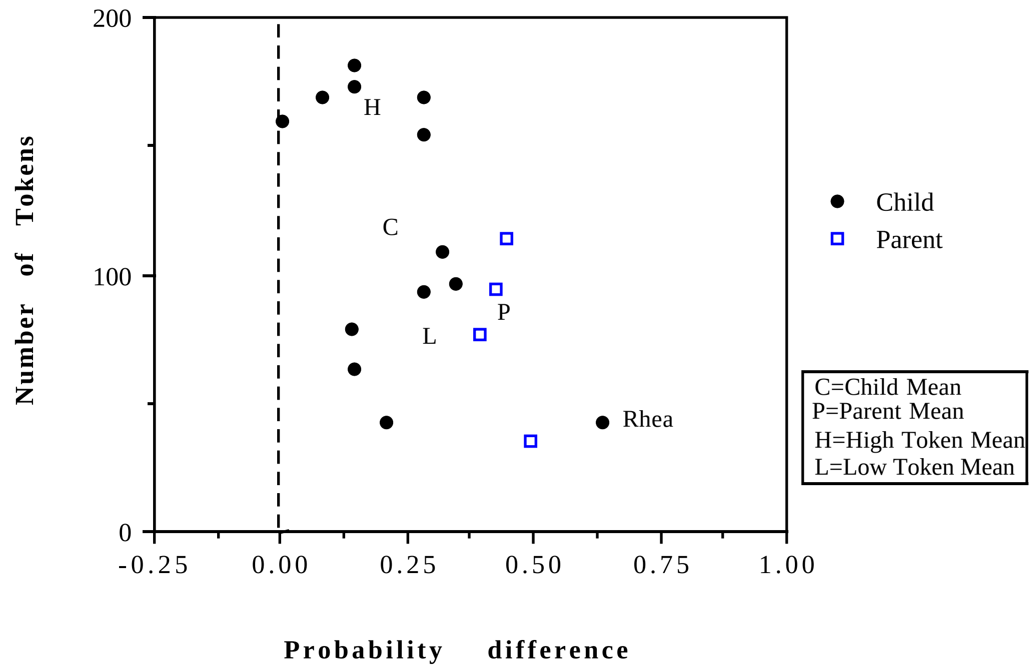
---

[16]The standard deviations in this table are quite large. Please note, however, that, as will be seen in the following discussion of the probabilities for (-t,d) deletion, it is the relationship between the percentages which is important not the size of the percentages themselves. Therefore, it does not detract from the results that the range of percentages in each category is wide. In addition, the the variance in the percentages is exaggerated by the standard deviation measure due to the large differences in the total numbers of tokens for each child.

themselves, but on the relationship between these probabilities. That is, one of the important findings in the effect of following segment on child speakers is that deletion of alveolar stops followed by consonants occurs significantly more often than that of alveolar stops followed by vowels or pauses. Therefore, the choice was made to show individual results by subtracting the probability of deletion for one category of following segment from the probability of deletion for another category of following segment. For example, for Figure 7.2, the probability of deletion of following vowel was subtracted from the probability of deletion of following consonant for each subject. A cluster around a given difference indicates that most or all of the subjects showed a similar difference of effect.

Figure 7.2 shows the difference between the effects of following consonant (obstruent and liquid) and vowel for adults and children. The subjects' probabilities cluster at a difference of .12 to .5 with the adults on the higher end of the range. Just as the results shown in Figure 7.1 show a difference between consonant and vowel for the group as a whole, so do the differences in individual results which cluster around the same probability range. Most important is the fact that for no subject is there a negative difference between consonant and vowel which would indicate that the subject was more likely to delete (-t,d) preceding a vowel than one preceding a consonant -- the opposite deletion effect to that shown by the group results and predicted by the (-t,d) deletion literature on adults.

Figure 7.2: Differences between the effects of consonant and vowel in adults and children.
p(following cons.) - p(following vowel)

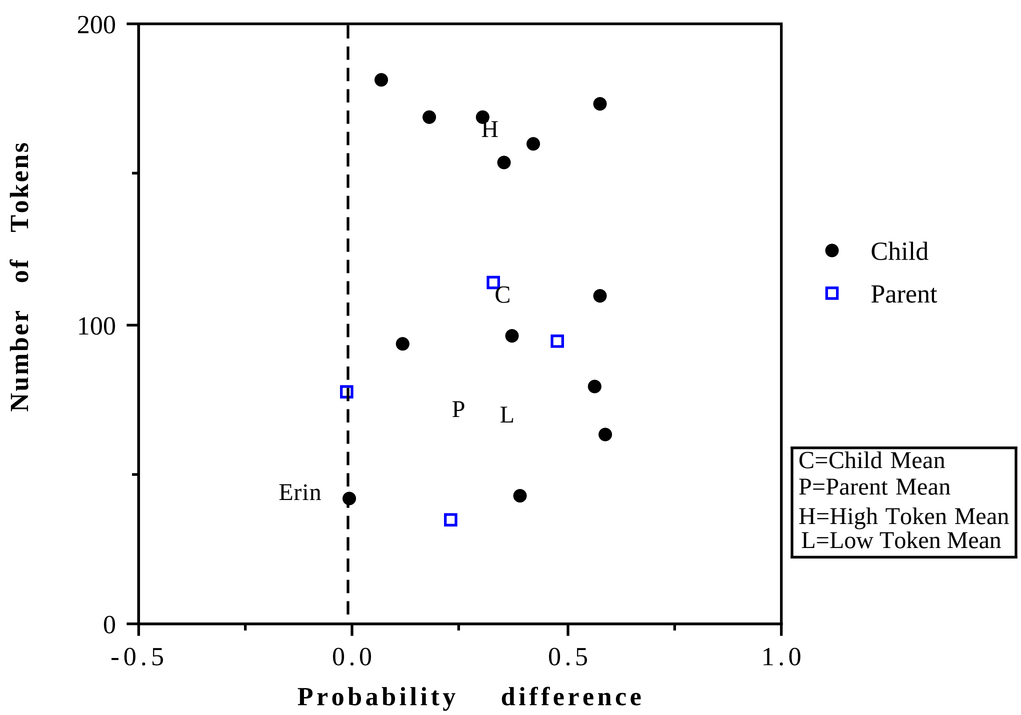


Along with the similarities among individual speakers in their uniform tendencies to delete (-t,d) before a consonant more often than before a vowel, there was a difference based on the age, or in this case, the generation of the speaker. As mentioned above, the adults probability differences were all at the higher end of the range. With only one exception, the probability differences of the children were at the lower end of the range. In other words, the adults were even more likely to delete (-t,d) before consonants than before vowels than were the children. The means of the parents' and children's probability differences appear in Figure 7.2, and there is a significant difference between them. ($mean_{children}$= .232, $mean_{parents}$=.439, t=2.554, $p<.05$, df=15.) It appears that, for this portion of the following segment constraint, the children, who are in the early phases of learning the constraints on (-t,d) deletion, are demonstrating a lesser difference than the adults, who have already learned them.

In looking at Figure 7.2, it becomes apparent that six of the children have appreciably more tokens (over 150) than the other seven (109 or fewer). The question that presents itself is whether or not this difference in the amounts of data corresponds to a difference in the results. In the case of the differences between the effects of following consonant and following vowel, it does not. The means of the probability differences for the High Token group and the Low Token group are shown in the figure, and even with the inclusion of Rhea's data, they are not significantly different ($mean_{high\ token\ group}$=.157, $mean_{low\ token\ group}$=.296, t=1.712, df=11). The difference in the variance around the means was also not significant (F=2.4, df=5,6.) In other words, although Rhea's probability difference diverges from that of the other children, this difference is not enough to affect the central tendencies of the two groups or the dispersion of data around the means. The children appear in this instance to present as a unified group of speakers showing a similar difference in effect of following consonant and vowel. The significant difference is between the children and adults in that, for the adults, the difference in the effects is greater than for the children.

The difference between the effects of consonant and pause is similar to that between the effects of consonant and vowel, as shown in Figure 7.3. The results of the subtraction of following pause from following consonant fall out between 0 and .5 for both adults and children. For the most part, the difference between the probabilities is consistent regardless of the number of tokens. Although the expected difference between consonant and pause is a positive one as shown by the group results, in this case, two subjects do show very small a negative difference. One adult speaker shows a difference of -.013, and one child a difference of -.006. Both of these speakers have total numbers of tokens below 100, on the lower end of the range for the group. The child (Erin), in fact, has a total N of under 50.

Figure 7.3: Differences between the effects of consonant and pause for adults and children p(foll. cons.) - p(foll. pause)



As with the previous figure, group means are shown on Figure 7.3. In this case, however, there is no significant difference between the means for the children and their parents in the differences in deletion between following consonant and pause ($mean_{children}$=.347, $mean_{parents}$=.254, t=.790, df=15.) Also, there is no significant difference in the means for the high and low token groups, again, indicating that differences in the amounts of data did not affect the primary results in this analysis ($mean_{high\ token\ group}$=.317, $mean_{low\ token\ group}$=.373, t=.474, df=11.) The differences in the variances are also not significant in either case (F=1.01, df=3, 12 for the children and parents; F=.571, df=6, 5 for the high and low token groups.)

We can say, then, that the results of the individual analysis of the effect of following segment fit that of the group analysis very well. The pattern shown by the group is not just a result of averaging a number of different individual figures but rather the result of pooling of like individual linguistic behaviors. In other words, the children in this study have acquired the following segment constraint on (-t,d) deletion. The children's (-t,d) deletion was most like that of their parents in the effect of following pause.

These findings are very important in examining the question of whether or not the acquisition of (-t,d) deletion can be considered rule learning at all or merely the influence of a performance factor on the production of consonant cluster. While some of the factors, for example following obstruent vs. following vowel, might be argued to be natural constraints based on ease of articulation, this is clearly not the case for following pause. Not only is there no articulatory explanation for this effect, but, as was noted in Guy (1980), it also varies by geographical area, with one rule for following pause in Philadelphia and a different one in New York. Although there is no data for New York children, these results show that these Philadelphia children do take their dialect of origin quite seriously. By the age of 3, they are indeed Philadelphia speakers like their parents. These data on following pause indicate that it is a socially learned dialect that is being acquired rather than a universal constraint being applied.

7.2 Effect of grammatical form on (-t,d) deletion

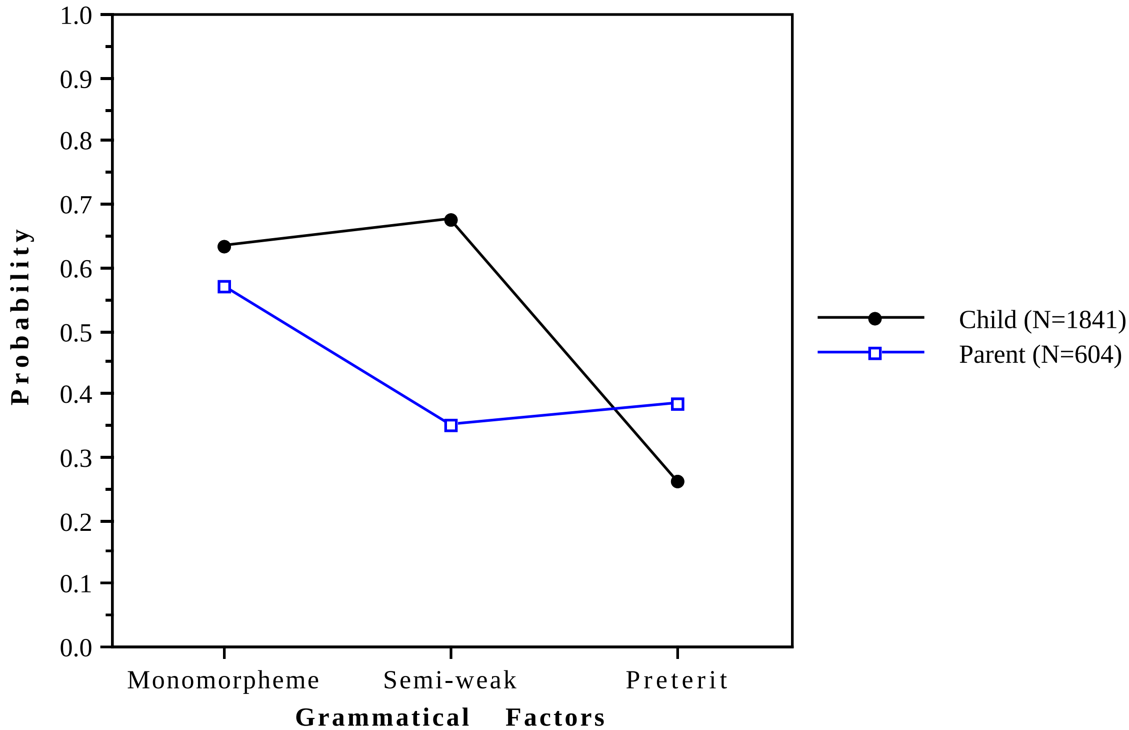
The results of the analysis of grammatical form are crucial to the second question addressed by this study, that of whether the children are learning rules or surface forms. Evidence that the children are, in fact, learning rules comes not so much from where the children follow the adult pattern but from where they differ. Guy (1980) noted that in the adults he studied, deletion was most likely in monomorphemic words, less likely in semi-weak past tense verbs, and least likely in weak past tense forms. As seen in Figure 7.4, the adults in this study showed the same difference in deletion between monomorpheme and semi-weak verb, but an insignificant difference in the rate of deletion between semi-weak verbs and weak past tense forms.[17] The semi-weak and weak past tense forms are included separately in the graph for comparative purposes only.

---

[17]Although tokens of (-t,d) which are contractions with the form (_n't) were included in the data, they were omitted from the analysis of grammatical form for two reasons. The first is that in previous adult research on the subject, contraction tokens were either not collected or were also omitted from the grammatical form analysis. Secondly and more importantly, for the children and adults in this study, there was wide individual variation, both in total deletion and in deletion in relation to other categories, in the contraction category.

The children, too, were more likely to delete /t/ and /d/ in monomorphemic words than in weak past tense forms. The similarities between the findings for adults and children end, however, with the semi-weak verbs. In this case, the children were as likely to delete in this category as in the monomorphemic category. In other words, there was no significant difference between the forms, and the children appeared to be making a similar analysis of these two categories as reflected in their patterns of deletion. These findings differ somewhat from those of Guy and Boyd (1990), who were working with considerably fewer tokens of child data, and, consequently, could only analyze the children as a group. They found that deletion in the semi-weak category was categorical for the youngest children and concluded that the /t/'s and /d/'s were not there at all for them. In the present study, however, deletion in the semi-weak category was not categorical, rather it was favored equally to the monomorphemic category. The semi-weak final segments appear to be present for these children but are analyzed as monomorphemes, rather than as productive past tense markers, as they are for most adults.

Figure 7.4: Effect of grammatical form on (-t,d) deletion.



As in the case of following segment, individual results were very consistent with those of the group as a whole. In Table 7.2, the percentages of deletion for the children are listed. In the monomorpheme category, the range of deletion is from 19 to 50%. The weak past tense and past participle forms were combined since they are similar both grammatically and in their probabilities of (-t,d) deletion. For participial and weak past tense forms, the range is from 5 to 25% with only one child (Mira) deleting over 18% of the time.

Table 7.2: Effect of grammatical factors on child subjects:

Name	Age	MLU	Mono. N	Mono. %	Semi-w. N	Semi-w. %	Part./Past N	Part./Past %	Total N
Cindy	3-3	3.43	72	40	6	83	55	11	133
Diane	4-6	4.12	32	38	4	25	16	6	52
Mike	3-4	4.52	35	26	0	-	37	8	72
Marie	3-2	4.76	47	19	0	-	24	13	71
Micky	3-5	4.78	51	25	2	0	38	5	91
Evan	3-10	5.01	49	39	1	100	20	15	70
Rhea	3-11	5.17	18	50	1	0	10	10	29
Erin	3-3	5.33	20	50	0	-	14	14	34
Jeanie	3-9	5.38	81	42	8	50	38	16	127
Zak	4-11	5.53	61	30	7	71	58	16	126
Mira	4-3	5.59	44	43	6	33	20	25	70
Jenny	3-11	5.82	86	30	8	38	41	15	135
Kent	4-6	5.82	28	39	4	25	11	18	43
Danny	4-10	5.90	57	46	10	60	57	16	124
Shelly	4-9	6.29	72	46	3	33	60	17	135
Gia	3-11	6.31	59	47	3	33	38	16	100
Mean			38.13		34.44		13.81		88.25
Standard Deviation			9.47		31.64		4.97		38.09

For this factor group, all 16 of the children provided enough data for individual analyses when, as in Table 7.2 above, participles

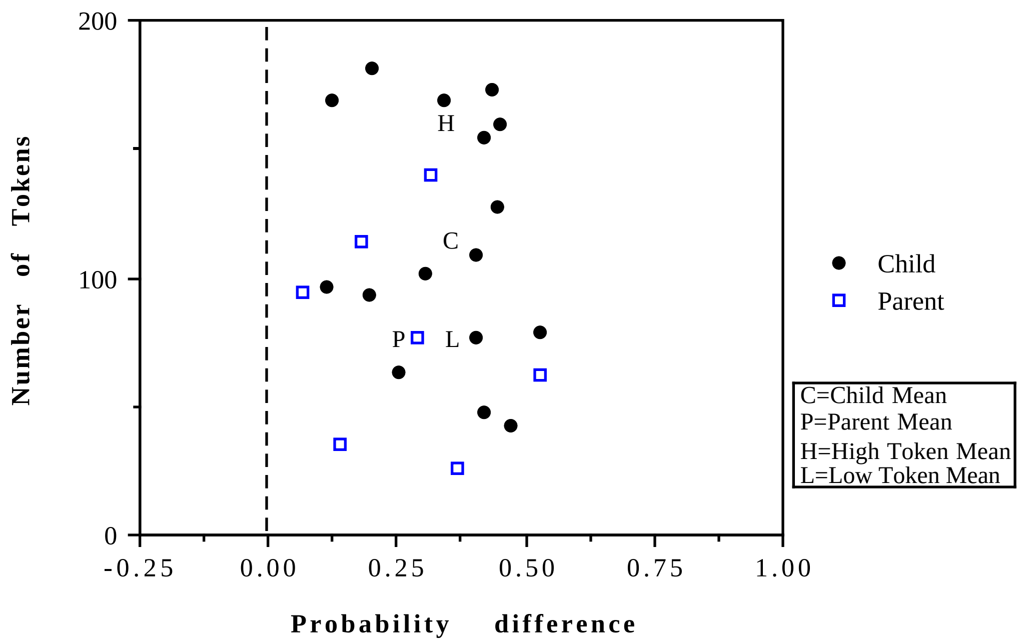
---

[18]The total N's for this table and for that of Table 7.1 are different because those tokens which are contractions were not included in this table. (See footnote #17.)

[19]It would have been helpful to the analysis to have been able to separate weak past tense verb forms and participles, since participles mark past redundantly. Unfortunately, as will be disussed further in section 7.9, there were far too few of these tokens, particularly from the Low MLU group, to have analyzed them separately. Participles, like the weak past tense verbs, seldom demonstrated (-t,d) deletion. However, it would have made a stronger argument against the children's participating in a universal tendency to preserve meaning to have been able to show the children's similar behavior in regard to these two forms in the variable rule analysis.

and weak past tense verbs were combined. The results, again, are very consistent. Figures 7.5 and 7.6 show the individual results for the effect of grammatical form. In Figure 7.5, the difference between the probabilities for deletion of monomorpheme and *-ed* forms (participles and weak past tense verbs) cluster at just below .25 to .5 for both adults and children with no instances of a negative difference. In other words, all of the children, like the adults, deleted (-t,d) more often in monomorphemic words than in *-ed* verb forms.

Figure 7.5: Difference between the effects of monomorpheme and -ed verb forms. p(monomorpheme) - p(-ed forms)

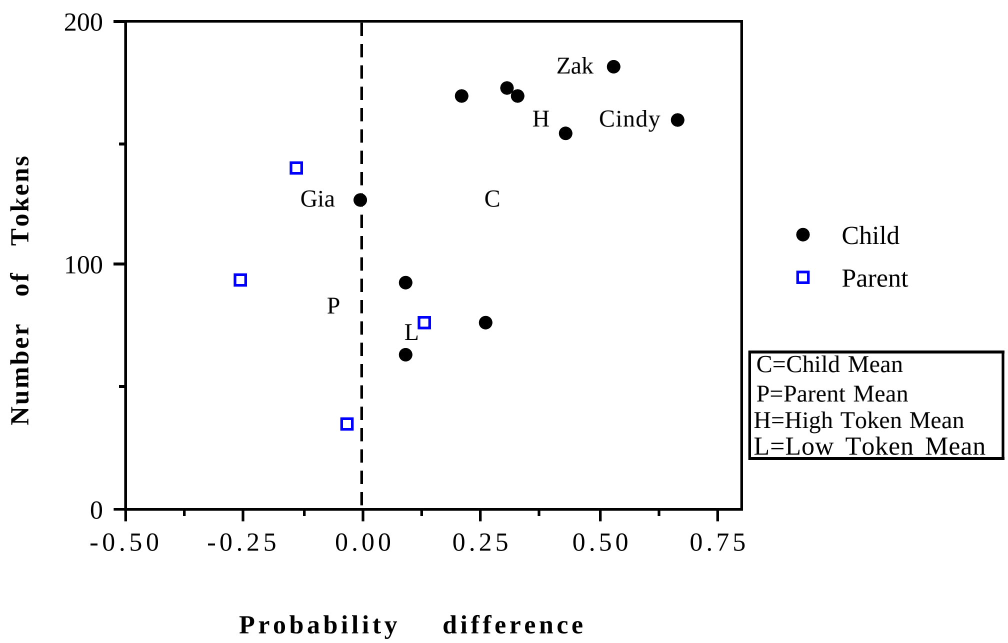


There was no significant difference between the means of the children and the parents in the effects of monomorphemic and *-ed* words ($mean_{children}$=.344, $mean_{parents}$=.269, t=1.205, df=21.) With the addition of the three children in the grammatical analysis, who were not included in the following segment analysis, the obvious division into high and low token groups disappeared. However, the children were divided along the same guidelines (Low Token Group: N=109 and below; High Token Group: N>109) in order to ascertain whether or not there were sufficient differences in the results of the analyses of the two groups to alter their means significantly. Again, these differences were not significant ($mean_{high\ token\ group}=.345$, $mean_{low\ token\ group}=.343$, t=.017, df=14), nor were the differences in variance statistically significant (F=.708, df=6,15 for parents and children; F=.944, df=8,6 for the high and low token groups.)

Figures 7.6 and 7.6a show a radically different picture from that of figure 7.5. In the case of the difference between semi-weak verbs and *-ed* forms, the children, with one exception, range from a difference of .1 to .6 while the adults are, with one exception, showing a negative difference. There is very little overlap between the adults and children indicating that the children are indeed analyzing the semi-weak forms much differently than are the adults. Only one child, Gia, showed a slight negative difference between the semi-weak and regular past tense verbs. Her (-t,d) deletion pattern does, on the surface, look very adult-like, but, since she only produced three semi-weak tokens, it is impossible to say at what rate she would delete (-t,d) in the semi-weak category if more data were available.

On the other hand, there are two children, Zak and Cindy, on the upper edge of this range. In these two cases, one might question, as did Guy and Boyd in their study, whether or not the (-t,d) forms are there at all for these children. If there, in fact, is a point at which children do not perceive the final (-t,d)'s in semi-weak verbs, Cindy would be among the most likely of the children to be at that point. At 3 years 3 months, she was one of the youngest children and had the lowest MLU. She was also a very talkative and cooperative child who attended Kids' Land five days per week, which accounts for the relatively large amount of data collected from her. Zak, on the other hand, was in the middle range in both age and MLU, and it is not immediately clear why his deletion of semi-weak verbs was as high as it was. However, neither child showed categorical deletion (Zak deleted in 5 cases out of 7, and Cindy deleted (-t,d) 5 times out of 6), and, again, with individual token amounts of this size, it is impossible to draw conclusions with any degree of certainty. The possibility exists that there is a stage at which children do not analyze semi-weak verbs as having final stops, but the fact that none of the children with more than one semi-weak verb token categorically deleted (-t,d) argues against it. In any event, if it is the case that a pre-awareness stage exists for final stops in semi-weak verbs, it appears to occur much earlier than noted by Guy and Boyd.

Figure 7.6: Difference between the effects of semi-weak verbs and -ed verbs. p(semi-weak) - p(-ed verbs)



The means of the parents' and children's differences between semi-weak and *-ed* verbs supports the finding that the adults and children are analyzing these forms differently. The difference between the mean for the parents (-.074) and that for the children (.290) is significant at the .01 level (t=3.087, df=12). The difference between the variances between the two groups is not significant, however (F=1.630, df=9, 3), indicating no difference in their dispersion around the mean.

As seen with the differences between following consonant and vowel, the total numbers of tokens appear to fall into two groups -- a Low Token Group at 127 and below and a High Token Group at 154 and above. The difference in the means for these groups is significant at the .05 level ($mean_{high\ token\ group}$=.411, $mean_{low\ token}$

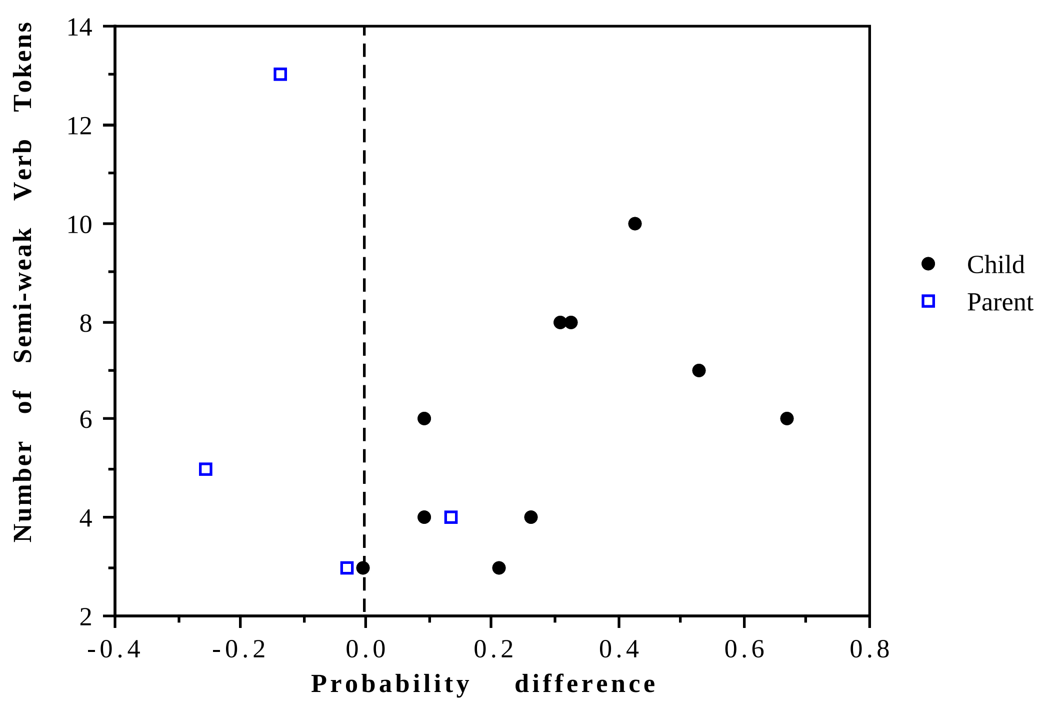
$_{group}$=.109, t=3.170, df=8). The difference in the variances for the high and low token groups is not significant (F=2.333, df=5, 3). The difference in the means is a puzzling finding in that it appears to be the opposite of what one might expect in that the speakers with lower numbers of tokens seem to be deleting (-t,d) more like the adults than those with high numbers of tokens. First, as will be illustrated further in the discussion of Figure 7.6a to follow, it is necessary to point out that we are working with very small numbers of semi-weak verb tokens, a fact which may, in itself, be the cause of this otherwise odd finding. In addition, however, the speakers in this low token group are there, in most cases, not because their language abilities are less mature than those of the other children but for completely different reasons. In fact, Gia, although in the middle of the age range, has the highest mean length of utterance.[20] Mira and Kent, also in this low token group are among the oldest children and have MLU's in the higher end of the range but were not tape recorded as often as the other children due to their infrequent attendance at the day care center. Only Diane, at age 4 years 6 months, is in both the low token group and the low MLU group, and the fact that she was one of the quietest and most difficult to engage in conversation of the children could well have contributed to her placement in these groups. In summary, although it is interesting to speculate on the subject of the difference between the means of these two token groups, the low quantity of semi-weak verb data and the complete lack of parallelism between membership in the low

[20]Please see chapter 6 for a more detailed discussion of MLU as it relates of age and language level.

token group and low age, MLU, or linguistic level make it impossible to answer the questions posed by this finding.

As a way of further illustrating the analysis of the semi-weak verbs in conjunction with the number of tokens in this category, Figure 7.6a plots the probability differences between the effect of semi-weak verbs and that of *-ed* verbs against the number of semi-weak tokens rather than total number of tokens. Even though the number of semi-weak verb tokens is small for both the children and the adults, the results show consistently that for the children the distinction is between the monomorphemes and semi-weak verbs versus the weak past tense forms, whereas for the adults, it is between the semi-weak and weak past tense forms versus the monomorphemes.

Figure 7.6a: Difference between the effects of semi-weak verbs and -ed verbs plotted against number of semi-weak tokens

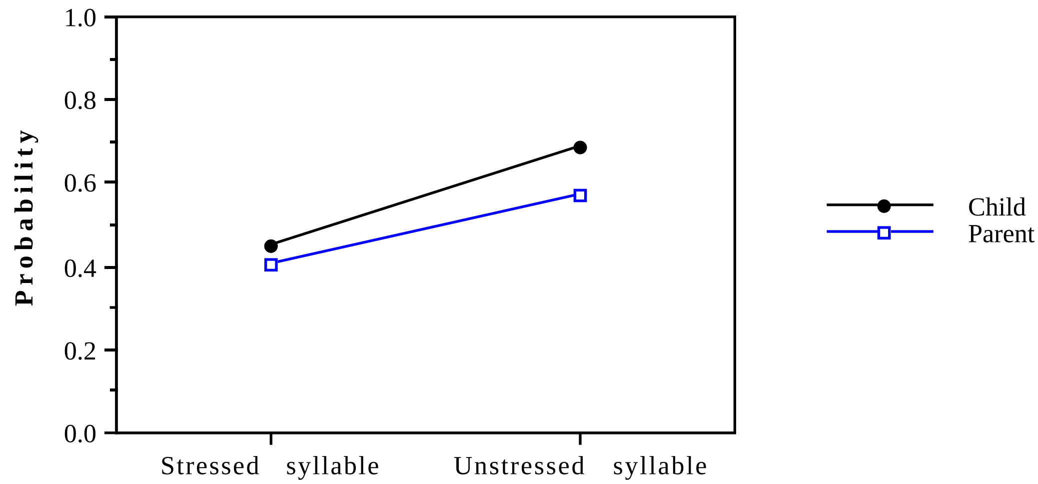


The fact that no child with more than one semi-weak verb token deleted (-t,d) categorically for that form indicates that these segments are real for the majority, if not for all, of the children. In addition, the fact that children do not duplicate the adult pattern for semi-weak verbs precludes any suggestion that they are simply imitating the surface forms produced by the adults from whom they are learning language. Instead, the effect of grammatical form on (-t,d) deletion in children is radically different from that in adults. This analysis of the semi-weak verbs as similar to monomorphemes in contrast to the adult analysis also has implications for the question of whether the children might be acting out of a natural tendency to preserve meaning. If this were so, the children, again, would have the deletion patterns of their parents exactly in line with the universal constraint, rather than deviating from it as they actually do. This deviation from their the adult pattern suggests that they are forming analyses of their own as part of a rule-learning process, not conforming to universal tendencies.

7.3 Effect of syllable stress on (-t,d) deletion

Fasold (1972), Wolfram (1972), and Labov (1989a) found that unstressed syllables were more likely to undergo (-t,d) deletion than stressed syllables. Syllable stress was also found to affect the presence or absence of (-t,d) deletion in the present study. As shown in Figure 7.7, both adults and children were more likely to delete /t/ or /d/ in an unstressed syllable than in a stressed one. While the difference is not significant for the adults, it is a significant difference for the children. ($p<.01$)

Figure 7.7: Effect of syllable stress on (-t,d) deletion.

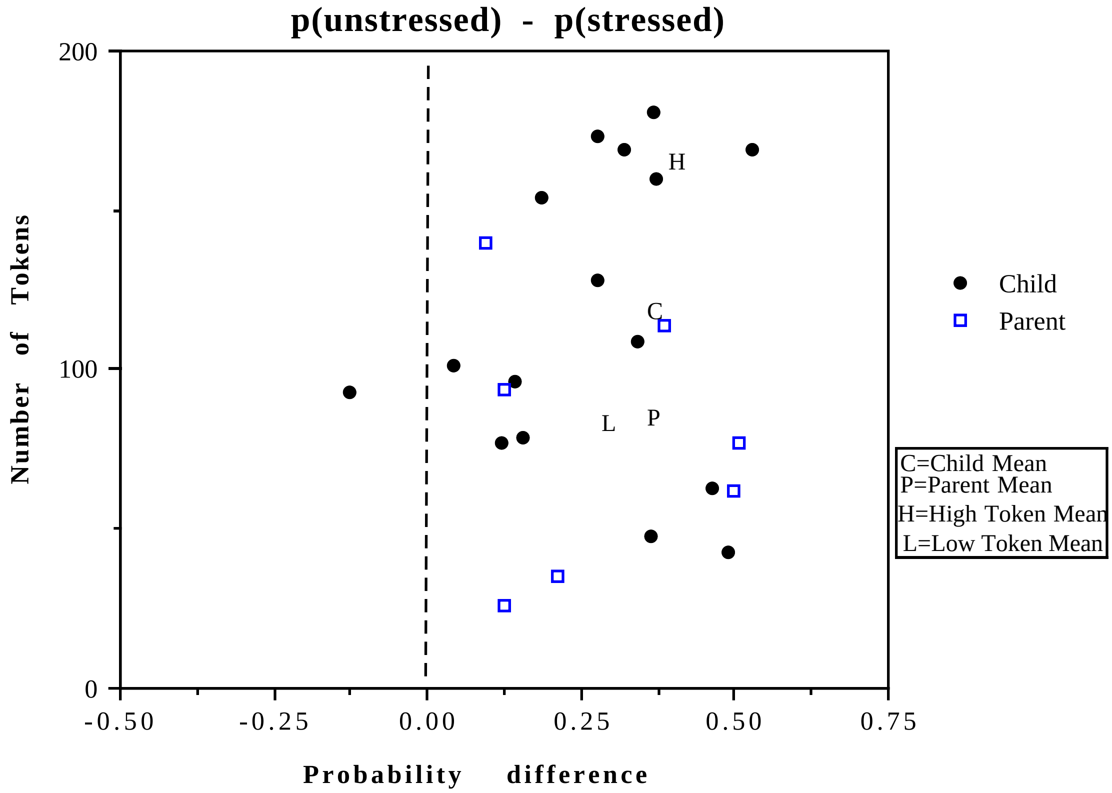


As in the cases of following segment and grammatical form, the individual data are quite consistent. Figure 7.8 demonstrates that all but one subject show the predicted finding that the presence of a (-t,d) cluster in an unstressed syllable favors deletion more than the presence of a (-t,d) cluster in a stressed syllable.[21] In addition, there is no significant difference between the means of the parents and children ($mean_{children}$=.270, $mean_{parents}$=.278, t=.104, df=21) or between the means of the high and low token groups ($mean_{high\ token\ group}$=.332, $mean_{low\ token\ group}$=.222, t=1.283, df=14.) Finally, the variances in both instances do not differ significantly (F=.939, df=6.15 for the parents and children; F=.262, df=8,6 for the high and low token groups.)

---

[21]In this case, the exception is Mira, a child in the Low Token Group due to her infrequent attendance at the center. She was otherwise indistinguishable from the other children, in that her speech resembled that of the group for the other factors examined in the analyses of both (ing) and (-t,d).

Figure 7.8: Difference between the effects of stressed and unstressed syllables on (-t,d) deletion.



7.4 Effect of the presence of a preceding third consonant on (-t,d) deletion

Labov (1989a) noted that the 7-year-old in his study had an even stronger tendency than the adults to favor deletion in the presence of a (-t,d) cluster preceded by a third consonant "as in *next* and *wouldn't*". Guy (1980) noted a similar tendency but did not examine it quantitatively using instead a measure of articulatory complexity. In the current study the presence of a preceding consonant as in *asked* or *next* did not significantly affect deletion. The presence of a preceding consonant and syllable boundary, as in the contractions *couldn't* and *wouldn't* and past tense verbs such as *struggled*, did favor deletion. (p>.01) There were several difficulties with the categories formed for this analysis, however, and the result is perhaps more questions than answers.

The first and most straightforward of the difficulties encountered in this analysis was the lack of single syllable tokens in which the (-t,d) cluster was preceded by a consonant. There were only 23 instances of this combination out of a total of 1841 (-t,d) tokens. It is difficult, if not impossible, to draw strong conclusions from this small amount of data.

Secondly, the preceding consonant tokens which included a syllable boundary consisted almost exclusively of the contractions *wasn't, couldn't, wouldn't,* and *shouldn't*.[22] Although, as discussed in footnote #14 in this chapter, the children were notably inconsistent in their deletion of contractions as a whole, and these forms were eliminated from the analyses of the grammatical factors affecting (-t.d) deletion. However, the children did appear consistently to delete (-t,d) in two-syllable contractions. The result is an interaction between the factor groups. It is not known whether it is the grammatical form, the syllable boundary or both which contribute to greater deletion in these tokens.

Thirdly, because of the neutralization effect which takes place in words having the form of a (-t,d) token preceded by an /n/ and followed by a vowel, these tokens were eliminated from the study. This means that the remaining two-syllable contractions were

---

[22]There were, in fact, three tokens with a syllable boundary and a (-t,d) cluster with a preceding consonant -- two instances of *snuggled* by the same speaker and one of *scrambled*.

followed, in most cases, by a consonant, and in a smaller number of cases, by a pause. Since, as discussed Section 7.1, we know that a following consonant favors deletion, there is again interaction between these two factor groups making it impossible to distinguish which factors are affecting deletion in the face of a limited variety of data in this area.

Finally, the factor group of syllable stress comes into play in this analysis, since, in those words with a syllable boundary, as noted above, the second syllable with the (-t,d) cluster is unstressed. Although, as discussed in chapter 4, syllable stress is not a strong factor affecting (-t,d) deletion, the results of this study, as well as others, have shown that consonant clusters are more likely to be reduced in unstressed syllables than in stressed syllables. Figure 7.9 shows that there are examples of words which are not contractions with and without preceding consonants in both stressed and unstressed syllables. As previously noted, however, instances of unstressed syllables containing a cluster with a preceding consonant were exceedingly rare. At the same time, the categories of preceding segment and syllable stress for contractions are mutually exclusive. That is, there are no possible instances of a contraction without a preceding consonant in an unstressed syllable or of a contraction with a preceding consonant in a stressed syllable. Once again, it is impossible to determine whether the unstressed syllables in the contractions or the presence of the third consonant is more influential in promoting deletion in these forms.

Figure 7.9 Interaction of preceding consonant and syllable stress in (-t,d) deletion.

Others

	stressed	unstressed
c(-t,d)	nest	opened
cc(-t,d)	next	snuggled

Contractions

	stressed	unstressed
c(-t,d)	can't	0
cc(-t,d)	0	couldn't

The preceding segment question continues to be unanswered with regard to young children, and further research is needed to shed light on this issue. The difficulties encountered by this analysis do, however, point to an additional dilemma of which the researcher must be aware when doing work on variation in child language. While children, even at the age of three and younger, can be rich sources of data for the sociolinguist as well as the psycholinguist, it is the case that their discourse tends to be more repetitive than that of adults. The fact that toys and props are often used to foster conversation can also add to this difficulty much as would bringing up the same topic of conversation over and over to an adult. Therefore, it is necessary when examining the results of an analysis to make sure that the data are not coming from a very small number of words. In the case of preceding segment, this is exactly what happened, and, therefore, the results merely indicate the need for further research but cannot provide answers.

7.5 Effect of style on (-t,d) deletion

Only the children's data was analyzed for differences in audience, since the adults were talking almost exclusively to me during the interviews. It was found that there were no significant differences for the children in terms of addressee. In other words, speaking to an adult, another child, or an inanimate "character" did not affect deletion.

Results were also not significant for the analysis of conversational style. A more informal, narrative style has been found by others to favor deletion in adults, but, as previously noted, equivalent style ratings could not be made for children. The children told very few narratives while being interviewed. The closest they came to it was during a pretend book reading activity, one that usually brought out a more formal "reading intonational" style as reported by Scollon and Scollon (1981). Style categories were examined based on the different activities the children engaged in, but these different categories did not result in different linguistic styles. The possibility that this method of interviewing and analysis did not capture the children's stylistic repertoires cannot be ruled out. Nevertheless, the findings of the analysis suggest that stylistic constraints on (-t,d) deletion, unlike the grammatical and phonological ones, have not yet begun to be acquired by the age of 3 and 4.

7.6 Effect of gender differences on (-t,d) deletion

Girls were more likely to delete /t/ and /d/ in consonant clusters than boys. These findings were significant at the .01 level (girls=.530; boys=.449). As shown in Table 7.3, the boys and girls showed similar patterns of deletion in both the effect of following sound and grammatical form. The girls, however, in most cases simply deleted (-t,d) more often.[23] These findings are particularly

[23]While, as stated above, the overall difference in probabilities of deletion

interesting because they are the opposite of that which one might expect given the results of previous adult studies.

Table 7.3: Variable weights for (-t,d) deletion for girls and boys.

Following Segment

Segment	*Girls (N=1156)*	*Boys (N=685)*
Obstruent	.732	.692
Liquid/Glide	.677	.514
Vowel	.549	.458
Pause	.376	.319

Grammatical Form

Form	*Girls*	*Boys*
Monomorpheme	.644	.581
Semi-weak Verb	.597	.696
-ed Verb	.227	.252

The adults in this study could not be used for comparison since the vast majority of the data came from women. Most adult studies of (-t,d) deletion have not included sex as a factor group, but both Wolfram (1969) and Neu (1980) found men to be more likely than women to delete (-t,d). Also particularly relevant is the Fischer (1958) study of the alternation between [ɪn] and [iŋ] which found girls to be more likely than boys to produce the more conservative [iŋ] form as opposed to the less standard [ɪn]. These findings are not surprising since men have generally been found to be more likely to

---

between the boys and girls is statistically significant, the differences between them for each of the factors, as illustrated in Table 7.3, are not. In other words, for example, there is no significant difference in the probability of boys and girls to delete (-t,d) before obstruents or in monomorphemic forms. It is only when the data from the various factors are combined that the significant difference between the sexes is revealed.

produce fewer formal or more prestigious forms in situations of stable variation and in situations of language change from above the level of consciousness. Women, on the other hand, have, in most cases, been found to use more of the newer, incoming forms in situations of language change from below the conscious level. This apparent dichotomy has been the focus of recent discussion in the area of sex differences in language.

Eckert (1989) notes that the oppositional nature of sex differentiation does not capture the social relevance of these categories in situations of language change. In fact, she rejects the term sex differences, as connoting a biologically based classification of effects, in favor of gender differences, a social construction of sex. Her study of Detroit adolescents shows that gender can have a variety of effects on linguistic variables which go beyond the continuum of more or less conservative. Labov (1990) also underscores the importance of social factors in the interpretation of sex differences in language change. Specifically, he discusses the interaction between sex and social class and states that, while the two factors may be independent at the beginning of linguistic change, the interaction between them increases with social awareness of the change. Therefore, he found that women often lead in the early stages of linguistic change, and their relatively greater influence on the language learning of young children accelerates these changes further. As social awareness of these changes increases, however, women frequently reject these changes, resulting in greater linguistic conservatism.

Given these findings, one would expect for the results of the analysis of (-t,d) deletion in children to show one of two things in regard to sex differences. Since, as Labov points out, women are more likely to be the primary model for early language acquisition,[24] it stands to reason that both boys and girls would be equally exposed to the more conservative language behavior of women as regards (-t,d) deletion, and would be deleting the forms more-or-less equally. The second possibility would be that the boys might be sufficiently exposed to the male norm and would be learning to delete (-t,d) at a more frequent rate than the girls. In fact, as was previously stated, neither of these predictions is correct. The girls are more likely to delete (-t,d) in this case than the boys. To speculate further on why this might be so requires one to move from the realm of fact to that of interpretation. The following discussion will do exactly that by presenting possible explanations for the findings and observations of the children's play behavior during the interview sessions.

At first glance, it is tempting to speculate that the girls are learning the (-t,d) deletion rule more quickly than the boys. There are indications from psycholinguistic research that girls do develop some grammatical forms earlier than boys, and the situation may be the same for the (-t,d) deletion variable rule. (See Maccoby and Jacklin (1974) and Wolf and Gow (1986) for further discussion of sex differences and language acquisition. However, the similarities between the sexes in the acquisition of the constraints on (-t,d)

---

[24]The fact that the children in the current study attend either part- or full-time day care, does not change this observation. Consistent with the norm in day care situations, the staff at Kids' Land was comprised entirely of women with the exception of the director's husband who was intermittently present at the center.

deletion show that no difference in their ability to learn the phonological and grammatical constraints. Therefore, it may be that the girls, having acquired the rule earlier, are applying it more frequently.

Another possible explanation for these findings lies in the dichotomy that women have often been found to produce more innovative (or fewer conservative) forms than men in situations involving changes from below the level of consciousness, but more standard forms in cases of stable linguistic variation or in situations above the conscious level. It is sometimes thought that women's relative lack of power and/or status in society may be linked with this greater linguistic conservatism in stable situations.[25] It seems likely that at the age of 3 and 4, girls have not yet responded linguistically to these societal conditions and are, therefore, less conservative than boys even in stable situations. In other words, it may be the case that for 3- and 4-year-olds, there may be no difference between linguistic behavior in situations of language change and of stable variation.

A third possibility is that since these children are in the process of acquiring language, they are in a situation that is for them one of language change. The fact that the girls delete (-t,d) at a higher rate than the boys then becomes consistent with the results for adults in situations involving language change.

---

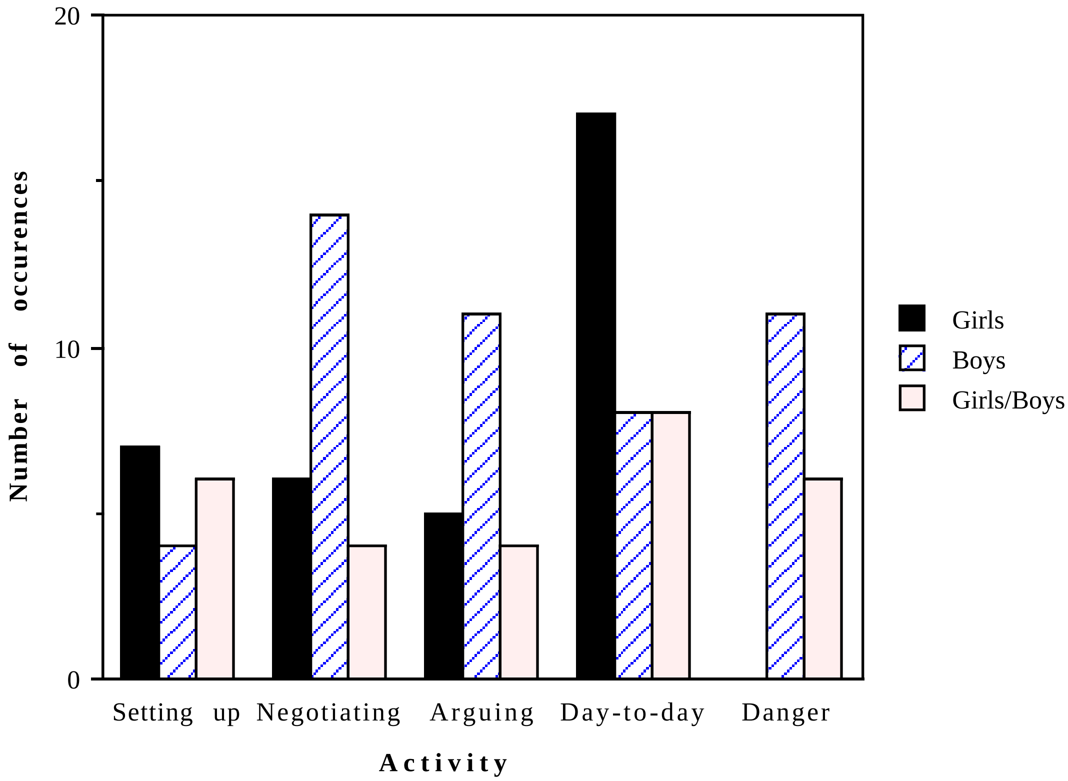
[25]This is an oversimplification of the interpretation of gender differences in studies of adult linguistic variation. Eckert (1989) points out that the basis of gender differentiation is rooted in culturally-based sex-role differences which, in turn, are linked to unequal allocations of power and economic opportunity.

None of this is to say that children have not learned any interactional behaviors linked to sex roles, however. Some general observations of the children's conversational topics as they played with each other indicate that some learning of this type has, in fact, taken place. As noted in Chapter 6, part of almost every interview session consisted of play with a Sesame Street toy which included a model of four rooms, toy furniture, and several characters which appear on the television show of the same name.[26] These play sessions lasted for approximately ten to twenty minutes, and the children were encouraged to interact with each other with as little interference from the interviewer as possible. The children were usually interviewed in groups of two, so five sessions of each of the three possible boy-girl combinations were randomly chosen for the observation. Play themes, initiated by the children, were listed for each session. The results are presented in Figure 7.10.

---

[26]The Sesame Street toy was felt to be appropriate for this type of observation since, although most of the characters are male, there are no sex stereotyped characters or objects included with the toy. That is, there are neither toy weapons nor babies or other parenting or housekeeping props, but only the characters and day-to-day household furniture, vehicles, outside props, etc.

Figure 7.10: Gender differences in occurences of conversational topics in play sessions.



All of the children participated in setting-up activities, usually at the beginning of each session, but also interspersed throughout many of the sessions. These consisted primarily of arranging the furniture within the rooms. All of the children also participated in negotiating and arguing about who got what character, piece of furniture, or car. Negotiating differed from arguing in that the latter was characterized by a rise in volume and pitch and a request for intervention from the interviewer. Both of these activities were interspersed throughout the sessions, and argument often directly followed negotiation. The boy dyads were more often found to be engaging in both of these activities than were the girl or boy-girl dyads.

The largest difference between the sexes, however, came in the choice of the play themes themselves. Both boys and girls participated in day-to-day household activities with the characters such as cooking, eating, going out to play, making telephone calls, cleaning, watching television, or going to work. The girls, however, were more than twice as likely to choose these themes as the boys or the boy-girl dyads. On the other hand, the girl dyads never focused their play on danger related themes. These consisted of car accidents, wind blowing over the house, fighting (by the characters), or injuries. This type of play was only in evidence during the sessions of the boy dyads or, to a lesser extent, the boy-girl dyads. Further, the danger play by the boy-girl dyads was, in every case, initiated by the boy.

This illustration suggests that these children have developed preferences for play which are consistent with sex-role stereotypes. That is, the girls appeared to be more likely to play at day-to-day activities, although both children participated in these activities. The boys, on the other hand, were the only ones who initiated the more dangerous or aggressive play activities. As noted, however, the results of the variable rule analysis show that these differences do not extend to (-t,d) deletion. In this case, the girls are more likely to exhibit the greater frequency of deletion more generally found in adult males than adult females.

It is impossible to say which, if any, of the explanations offered in interpreting this data is persuasive. However, it is clear that at the age of 3 or 4, these girls have begun the process of learning culturally based sex-role behaviors but have not yet learned their role as guardians of the conservative linguistic norm. These findings are clearly supportive of those of Eckert and Labov in that they argue against a biological explanation of sex differentiation in linguistic behavior. They also suggest that the children do form their own analyses of their language, and that the frequencies of their deletions are also their own, not copied from the adults around them.

7.7 Effect of age and MLU on (-t,d) deletion

The children ranged in age from 3 years, 2 months to 4 years, 11 months, but dividing the children into groups of ten 3-year-olds and six 4-year-olds did not yield significant differences in regard to (-t,d) deletion. As discussed in Chapter 6, the children were also divided into Low and High groups by mean length of utterance (MLU). The Low MLU group consisted of eight children with MLU's ranging from 3.43 to 5.33 morphemes per utterance. The High MLU group was made up of eight children with MLU's of 5.38 to 6.31 morphemes per utterance. This division, like the one by age, did not yield significant results in terms of overall probability of deletion. There are differences, however, in the types and numbers of grammatical forms produced by the two groups. The total number of (-t,d) tokens produced by the groups differed greatly in that the high MLU group produced 1129 (-t,d) clusters and the low MLU group 714.

More important than the differences in total numbers of tokens is the difference in the number of some of the more advanced grammatical forms.[27] The high MLU group produced significantly

[27]While these findings appear to be in contrast to those of Klee and Fitzgerald

more participles and semi-weak verbs than did the low MLU group as shown in Table 7.3. The difference between the groups in terms of their production of monomorphemes and weak past tense verbs was not significant. It appears clear that monomorphemes are neither syntactically nor morphologically complex forms, since they have no internal boundaries and can, and generally do, make up children's early one- and two-word utterances. One would expect them to occur frequently in these children's utterances, and they do in both MLU groups. On the other hand, participles require, in most cases, a syntactically complex frame including an auxiliary verb or modal (e.g. The food was cooked.). Neither group produced many of these forms, and the Low MLU group, who by definition have shorter utterances, produced proportionately even fewer than the High MLU group.

The differences in the groups' production of semi-weak and weak past tense verbs is not so clearcut, but a review of the discussion on the acquisition of past tense in Chapter 2 may help to clarify them. Brown (1973) and Ervin and Miller (1963) both note that irregular verbs are the first to appear in children's language. Kuczaj (1976), however, argues that children reliably acquire the weak past tense verbs considerably earlier than they do the irregular verbs. The children in the current study had to have acquired the weak past tense rule in order to be included in the analysis of (-t,d) deletion. There was no such criterion for irregular

---

(1985), discussed in Chapter 6, note that their analysis of grammatical acquisition was based on the *Language Assessment, Remediation and Screening Procedure (LARSP)* (Crystal, Fletcher, and Garman, 1976) which analyzes a variety of structures at the word, phrase, and clause level. The present study, however, looked only at a few specific morphemes.

verbs, however, since their acquisition was not crucial to an analysis of final stop clusters. Since the semi-weak verbs are a subset of irregular verbs, it is possible that the children with the lowest MLU's have not acquired all of the semi-weak verbs as consistently as have the children with higher MLU's.

Table 7.4: Differences in numbers of participles and semi-weak verbs produced by High and Low MLU groups (p<.01).

Group	Participle	Semi-Weak	Total Tokens
Low	11 (2%)	14 (2%)	714
High	47 (4%)	49 (4%)	1129

It appears that the two groups are deleting /t/ and /d/ at equivalent rates, but the Low MLU group is producing fewer of the participial and semi-weak forms, and the deletion for this group is, therefore, taking place in the monomorphemic and, to a lesser extent, in the weak past tense verb categories. As the children start producing these tokens, however, they appear to be simplifying them in a manner similar to the High MLU group. For example, of the eleven participle forms produced by the Low MLU group, only one contained a deletion. Of the fourteen semi-weak forms produced, however, seven had a (-t,d) deletion. Although the amounts of data for these forms are small, when taken in conjunction with the overall results for both groups, it seems that the children are learning the (-t,d) deletion rule simultaneously with the acquisition of the grammatical forms themselves.

7.8 Individual differences in (-t,d) deletion

Differences among the individual subjects as to the overall probability of (-t,d) deletion were not significant. This was true for both adults and children. As previously discussed in the analyses of the different factor groups, the individual subject findings were very like those for the entire group in terms of the following of the constraints on deletion. Similarly, in the findings for overall deletion, the subjects were deleting at rates that were not significantly different from each other.

7.9 (-t,d) deletion and past tense acquisition

Chapter 2 of this study contains a discussion of the connectionist and the rule-learning approaches to the study of children's overregularizations of weak past tense verbs. The question that is basic to the argument between the approaches is, what exactly are children learning as they learn to produce past tense verbs. The connectionist approach suggests that associations are being formed and strengthened by repetition. These associative networks are based on phonological similarity, and, possibly, semantic information. Grammatical category would be to a large extent irrelevant in this approach as would be rules and verb roots. The rule-learning approach to past tense formation is based on the supposition that rules are not only relevant, but basic to the acquisition. As previously noted, Kim, Pinker, Prince, and Prasada (1991) address this issue directly when they show that adults rely on grammatical categories in the past tense formation of novel verbs. They found that verbs with noun roots were judged better with regular past tenses, while verbs with verb roots were thought to be better with irregular past tenses.

While this study does not address the question of how children learn past tense directly, its findings are relevant to this argument. Most importantly, the results of this analysis show that children do learn variable rules. The dialectal specificity of the following segment constraint as well as the age or generational differences in the grammatical form constraint cannot be accounted for by linguistic universal, word-based probability matching, or phonologically based neural connections. Furthermore, just as Kim et al.'s findings showed that adults are sensitive to grammatical category and make use of it when forming the past tenses of novel verbs, the present findings on the acquisition of the grammatical form constraint demonstrate that children, too, share this sensitivity.[28] They not only learn this constraint from their caretakers as shown by their deletion in monomorphemes and weak past tense verbs, but they also demonstrate by their unique analysis of semi-weak verbs that they attach the probabilities to the abstract grammatical forms, not to individual words learned from those around them. Although, as stated above, these findings are not specifically on the subject of past tense formation, they are nonetheless useful. Unlike the work of Kim et al. which used adults as subjects, the present study looks directly at children who are in the process of learning language. The results show that the

---

[28]To a lesser extent because of its relative simplicity, the findings on the acquisition of the (ing) production rule discussed in the next chapter also support the importance of rule-learning and sensitivity to grammatical form in the acquisition of variation.

sensitivity to grammatical structure is not something which develops over time, nor is it a result of formal education. Rather, the use of grammatical structure in learning rules is a process that begins with the acquisition of the first grammatical forms.

7.10 Summary

The major findings discussed in this chapter can be summarized as follows:

1. The most important finding of this analysis is that the 3- and 4-year-old subjects in this study had to a major extent acquired the (-t,d) deletion rule and its grammatical and phonological constraints. They were also beginning to demonstrate social differences in (-t,d) deletion, at least in regard to sex differences.
2. The children had acquired both the constraint of following segment and that of syllable stress, in that for both they very closely mirrored the adult pattern. Of particular note was the children's acquisition of the dialectically based following pause constraint which suggests that they are in the process of acquiring a socially learned dialect of which the following segment constraint is a part.
3. By the age of three, the children were also well on their way to acquiring the grammatical constraint to (-t,d) deletion. Like their parents, they were significantly more likely to delete (-t,d) in monomorphemes than in *-ed* verbal forms. Their departure from the adult pattern of

decreased probability of deletion for semi-weak verbs, however, suggests that they are forming their own analysis of these verbs in which the /t/ or /d/ is not a semantically salient part of the word. The ability of these children to form this analysis which differs from that of their parents indicates that they are actually learning a deletion rule rather than imitating surface forms. This unique deletion pattern found in the analysis of the children's deletion pattern also indicates that, by forming their own analysis, they are not participating in a universal tendency to preserve meaning in speech.

4. While it is far more common in sociolinguistic literature for women to use more conservative, standard linguistic forms than men, in this study the girls were more likely to delete (-t,d) than the boys. While there is no clear explanation for these results at the present time, they do speak strongly against a biological basis for linguistic differences between males and females.

5. There were no significant differences between the children in (-t,d) deletion based on age or mean length of utterance (MLU). It should be noted, however, that the Low MLU group produced significantly fewer of the more complex semi-weak and participial forms than the High MLU group. When they did produce them, however, they deleted (-t,d) in a predictable fashion according to the deletion rule and its constraints. These findings suggest

that as the children acquire these forms, they simultaneously acquire the (-t,d) deletion rule.

6. The following table summarizes the more specific findings of this study by comparing them to the predicted findings from the literature on (-t,d) deletion in adults.

Table 7.5: Summary of results on (-t,d) deletion.

Constraint	Predicted Result	Actual Result
Following segment	cons>vowel>pause	cons>vowel>pause
Grammatical form	mono>semi-weak>weak	(mono=s-w)>weak
Syllable stress	unstressed>stressed	unstressed>stress
Audience	n/a	not significant
Style	informal>formal	not significant
Sex	males>females	females>males
Age/MLU	n/a	not significant
Individual differences	n/a	not significant

CHAPTER 8: (-t,d) DELETION AND LEXICAL PHONOLOGY

8.0 Introduction and literature review

Guy (1991) noted that in spite of the quantitative sophistication of variable rule analysis, there is little explanatory precision to the empirical studies in sociolinguistics. For example, traditional post hoc descriptions of the findings commonly reported for the effect of the grammatical constraint on (-t,d) deletion are either functional or structural. The functional description states that increased functional load carries a decreased likelihood of deletion. That is, the weak past tense -*ed* morpheme carries the past meaning, and, therefore, would be least likely to be deleted. On the other hand, the final stops on monomorphemic words carry no independent meaning, and would be most likely to demonstrate (-t,d) deletion. The semi-weak verbs, in which the final /t/ or /d/ is not a unique marker of past tense would have an intermediate status in terms of deletion.

The structural description relies on the morphological boundary which differs for the bimorphemic words, which have a /#/ boundary, the semi-weak verbs, which have a /+/ boundary, and the monomorphemic words, which have no internal boundary. Like increased functional load, the /#/ boundary most discourages deletion. The semi-weak verbs, with their /+/ boundary, are more likely to experience deletion. Finally, monomorphemic words are most likely to demonstrate (-t,d) deletion. Valid as these descriptions may or may not be, however, they do not make possible predictions for the probabilities of (-t,d) deletion in future studies. Instead, they simply offer descriptive devices after results are obtained.

Guy suggests an alternative explanation with predictive ability based on lexical phonology as discussed by Kiparsky (1982, 1985) and Mohanan (1986). In this approach, rather than differing by functional load or morphological boundary, the grammatical categories differ by the derivational level at which these forms acquire their final cluster. At each level of lexical derivation, the morphological processes alternate with the phonological processes. In addition, phonological rules may apply at any level. According to lexical phonology theory, boundaries, such as /+/ and /#/ discussed above, do not exist. Rather, bracketing marks morphemes, and bracket erasure occurs at each level, protecting the internal structure of the word from availability to rules later on. Therefore, since the (-t,d) deletion rule can apply at various levels, a weak past tense form might be exposed to the deletion rule only once, a semi-weak form twice (depending on how it is analyzed), and a monomorpheme three times. The result would be that weak past tense verbs would be less likely to undergo (-t,d) deletion than would semi-weak verbs, which would in turn be less likely to undergo deletion than monomorphemes. Another way of stating this, which more clearly demonstrates the mathematical predictions based on this hypothesis would be the following: Weak past tense verbs have a probability of retention based on the individual speaker's input rule application (Pr); semi-weak verbs probability of retention would be equal to Pr^2;

and monomorphemic words would have a probability of retention equal to Pr^3.

Guy tested this approach on the speech of eight subjects, aged 8 to 55. He found that his exponential explanation accounted for his results better than the more standard logistic one. The least accurate fit of the data was for the semi-weak verb class. Guy noted that there were fewer tokens in this class than in the others and that subjects of different ages appeared to analyzed this group of verbs differently. (See Guy and Boyd, 1990 and the discussion of this study in Chapters 3 and 4.)

Santa Ana (1991) confirmed Guy's lexical phonology model with data from 45 speakers of Chicano English in Los Angeles. He found, however, that his oldest speakers, those born before World War II, exhibited a high deletion rate for semi-weak verbs as well as monomorphemes. These results were puzzling since, according to the findings of Guy and Boyd (1990), it is these older speakers who would be most likely to have differentiated the semi-weak verbs as a separate morphological class. Santa Ana notes, however, that the group of speakers is unique in the community in that they often used Spanish for peer communication. It was found that if these speakers were considered to have two morphological classes -- weak verbs and monomorphemes combined with semi-weak verbs -- there was solid confirmation of the exponential model in the Chicano English speech community.

Bayley (1993) found similar results to those of Santa Ana in his study of San Antonio Tejano English. He combined the monomorphemic and semi-weak word classes and assumed that they were subject to three passes of the (-t,d) deletion rule. Following this assumption, he, too, found a strong confirmation of the variable lexical phonology model.

8.1 Results of the present analysis

The current study presents another challenge for Guy's exponential hypothesis since the data come from very young children who are in the process of acquiring the (-t,d) deletion rule as part of learning their first language. One might expect that the fit might be compromised by the facts that children's linguistic output in the preschool years is in a state of constant change as they acquire their native language and that the amount of data per child is, at times, less than optimal. Nevertheless, the data do provide support for the exponential model as will be shown below.

As discussed in the previous chapter, the probability of deletion in semi-weak verbs was not statistically distinguishable from that in monomorphemes. These results suggest that, for the children, as for Santa Ana's pre-World-War-II Chicano English speakers and Bayley's Tejano English speakers, there are only two classes of words undergoing deletion -- monomorphemes and semi-weak verbs in one class and weak past tense verbs in the other. Since there was found to be no significant difference between the semi-weak verbs and monomorphemes in the variable rule analysis of the children's data, they will be grouped in the following presentation of results.

Table 8.1 shows number and percentage of tokens which were retained as compared to that predicted to be retained. The differences between the predicted and observed retained tokens resulted in an error rate of 7%.

Table 8.1: Observed and expected frequencies for (-t,d) retention in 3- and 4-year-old children.

		Retention:					
Class	*Token count*	*% obs.*	*% pred.*	*Tokens obs.*	*Tokens pred.*	*% error*	*Token error*
M+S	875	61.7	P^3=.688	540	602.0	.07	62
P	478		P=.883	422	-----	-----	-----

Following Guy (1991), a best fit estimate of the probability of retention (P_r=.85593) using a chi square minimization technique resulted in a better fit of the data to the model as shown in Table 8.2. In this case, the error rate was less than 1% (.009) for the monomorphemes and semi-weak verbs combined and less than 3% (.026) for the weak past tense verbs. A chi square analysis revealed the results not to be significant at the .05 level, the desired result for a fit of model to data.

Table 8.2: Observed and expected frequencies for (-t,d) deletion and retention in 3- and 4-year-old children using a best-fit estimate of Pr.

	Total	*Observed*		*Predicted*		*Token*	*%*
Class	*Count*	*Del.*	*Ret.*	*Del.*	*Ret.*	*error*	*error*
M+S	875	335	540	326.3	548.7	8.7	.009
P	478	56	422	68.9	409.1	12.9	.026

$x^2=3.19$ 1 df
$p>.05$

As discussed in Chapters 7 and 9, it is critical to examine individual as well as pooled data. It is to be expected that pooled data, because of the considerably bigger N's will supply a better fit for the exponential hypothesis. In spite of the differing input rule applications of each speaker, however, their ratios of retention should be similar if the model is to be supported by this data. Table 8.3 lists the individual values of (-t,d) retention.

Table 8.3: (-t,d) retention by individual children.

	Token	*Obs.*	*%*	*Est.*	*Best*	*Exp.*	*Diff.*
Name	*count*	*ret.*	*ret.*	*Pr*	*Pr*	*ret.*	*(o-e)*
Cindy							
M+S	78	44	56.4	.826		45.74	-1.74
P	53	47	88.7	.887	.837	44.36	2.64
Diane							
M+S	36	23	63.9	.861		23.62	-.62
P	14	13	92.9	.929	.869	12.17	.83
Mike							
M+S	35	26	74.3	.906		26.29	-.29
P	36	33	91.7	.917	.909	32.72	.28

Marie							
M+S	47	38	80.9	.932		37.08	.92
P	24	21	87.5	.875	.924	22.18	-1.18
Micky							
M+S	53	40	75.5	.911		41.00	-1.00
P	38	36	94.7	.947	.918	34.88	1.12
Evan							
M+S	50	30	60.0	.843		30.06	-.06
P	20	17	85.0	.850	.844	16.88	.12
Rhea							
M+S	19	10	52.6	.807		10.40	-.40
P	9	8	88.9	.889	.818	7.36	.64
Erin							
M+S	20	10	50.0	.794		10.83	-.83
P	12	11	91.7	.917	.815	9.78	1.22
Jeanie							
M+S	89	51	57.3	.831		51.81	-.81
P	31	27	87.1	.871	.835	25.89	1.12
Zak							
M+S	68	45	66.2	.872		44.32	.68
P	53	45	84.9	.849	.867	45.95	-.95
Mira							
M+S	44	29	58.0	.834		27.94	1.03
P	20	15	75.0	.750	.824	16.48	-1.48
Jenny							
M+S	94	65	69.1	.884		66.49	-1.49
P	37	35	94.6	.946	.891	32.97	2.03
Kent							
M+S	32	20	62.5	.855		19.72	.28
P	11	9	81.8	.818	.851	9.36	1.64
Danny							
M+S	67	35	52.2	.805		36.27	-1.27
P	49	42	85.7	.857	.815	39.94	2.06
Shelly							
M+S	75	41	54.7	.818		42.88	-1.88
P	45	40	88.9	.889	.830	37.35	2.65
Gia							
M+S	62	33	53.2	.810		33.56	-.56
P	29	26	89.7	.897	.822	23.84	2.16

This table demonstrates that the individual data for the children also is a surprisingly close fit with the exponential model given the small amounts of data in some categories. Column 5 contains the estimated probabilities of retention for the monomorpheme and semi-weak verb classes. These two estimates should be as close to equivalent as possible for the best fit to the model. The differences between them range from .007 to .123. The greatest difference is seen in Erin's data which contains only 32 tokens in total. The mean difference is .055. Column 7 shows the expected retention rates for each category based on the minimum chi-square estimate of Pr. The final column reveals the difference between the expected and observed rates of retention. Fifteen of the total 32 expected rates of retention are within one token of the observed rate, 12 are within two tokens, and the final 5 are within three tokens.

The final table, Table 8.4, is a listing of the total chi-squares and significance measures for the individual children. The chi-squares are calculated using the expected and observed rates of retention and deletion in each of the two categories -- monomorpheme in combination with semi-weak verbs as one category and weak past tense verbs as the other. Each of the measures is based on one degree of freedom.

Table 8.4: Chi-squares and significance measures for individual children.

Speaker	*Total chi-square*	$p >$
Cindy	1.124	.20
Diane	.480	.30
Mike	.039	.80
Marie	.936	.30
Micky	.546	.30
Evan	.006	.90
Rhea	.339	.50
Erin	.961	.30
Jeanie	.319	.50
Zak	.178	.50
Mira	.841	.30
Jenny	1.262	.20
Kent	.103	.70
Danny	.672	.30
Shelly	1.298	.20
Gia	1.100	.20

Again, the desired result is a small chi-square which indicates a lack of a statistical significant difference between the expected and observed rates of retention. In this case, there is no chi-square over 1.298 and no significance measure in which p is less than .20.

8.3 Summary and Conclusions

The results of this analysis provide support for the exponential model from very young speech community members. Since the probabilities of deletion of monomorphemic words and semi-weak past tense verbs were found not to be significantly different in the variable rule analysis, these grammatical categories were combined. The result was an accuracy level of over 97% for the group data. The individual data was also found to be a good fit for the model in that in no case were the differences between the observed and predicted retained (-t,d) clusters found to be significant at above the .20 level.

When dealing with analyses of the language of very young children, it is often tempting to speculate on the possibility of linguistic universals. In this case, however, it must be remembered, that (-t,d) deletion and its constraints are an English language phenomenon. Therefore, the interpretation of results must also be language specific. Nevertheless, the speech of these 3- and 4-year-olds does provide independent confirmation for variable lexical phonology. Further, it appears, as was suggested by the variable rule analysis itself, that as children are learning variable rules and their constraints, their (-t,d) deletion patterns are consistent with theoretical predictions based, up until now, on the language of adult English speakers.

CHAPTER 9: RESULTS OF THE ANALYSIS ON (ing) PRODUCTION

9.0 Introduction

The following chapter will outline the results of the analysis of (ing) production. As noted in Chapter 5, there are fewer internal linguistic constraints on (ing) than there are on (-t,d) deletion. Phonological conditioning on (ing) production appears to be limited to regressive assimilation, and will not be examined in this analysis. On the other hand, the grammatical constraint on (ing) production will be examined in detail in the following section. External constraints on (ing) production will also be discussed, including those of effect of style, sex, age or linguistic level, and individual differences.

One of the questions to be examined in this analysis of (ing) production will be that of when children learn this alternation of [ɪn] and [iN] in comparison to their learning of the (ing) forms themselves. As stated in the preceding chapter, the results of the analysis of (-t,d) deletion acquisition by these children suggests that the variable rule is learned very early, possibly in conjunction with the learning of the past tense rule itself. Brown (1973) notes that the present progressive verb is one of the earliest inflectional morphemes to emerge in young children; therefore, it will be difficult to state with certainty whether the alternation between forms and the inflectional morpheme emerge simultaneously in 3- and 4-year-old children. The patterns of (ing) production for each child will be examined, however, for clues as to the process of acquisition in hopes of shedding light on this issue.

In addition, I will look at whether any findings on the acquisition of (ing) variation in these children can be explained by their copying the individual forms from their parents. The evidence for (-t,d) deletion presented in the preceding chapter indicates that this is not the case since the children do not demonstrate the same analysis of semi-weak verbs as their parents do. While there is no ambiguous form similar to the semi-weak verb in (ing) production that would enable us to state more definitively that children are not copying forms from their parents, examination of the individual tokens can provide useful clues in the examination of this issue.

Finally, the question of whether or not the children may be learning the velar and alveolar forms of (ing) as two separate morphemes will be discussed. In other words, it will be important to determine whether the children are performing a grammatical analysis in which the nouns contain an {iN} morpheme and the verbal forms an {ɪn} morpheme, or whether they are demonstrating rule-governed variation within a single morpheme. Again, the data will be examined for any evidence that may be brought to bear on this question.

9.1 Effect of grammatical form on (ing) production

As discussed in Chapter 5, grammatical constraints on (ing) production were discovered by a seminar in the Study of the Speech Community in the early 1980's at the University of Pennsylvania under the direction of William Labov. It was found that the [ɪn] was favored most in progressive verbs and participles/adjectives, less in gerunds, and least in nouns. This discovery was supported by the work of Houston (1985) using speakers from various parts of England. As can be seen in Figure 9.1, the adults in the present study were more likely to use the [ɪn] form in verbs and complements than in nouns and adjectives. Neither the difference between nouns and adjectives nor that between verbs and complements was significant, but the difference between the noun/adjective group and the verb/complement group was significant at the .01 level. The children in this study demonstrated (ing) production very much like that of their parents and also like results previously reported for adults. For the children, the [ɪn] form was most prevalent in verb and complements, less prevalent in verbal adjectives[29], and least prevalent in nouns.[30] The findings were significant at the .01 level for the children except for the difference between verbs and complements which was not statistically significant. Therefore, these two categories were

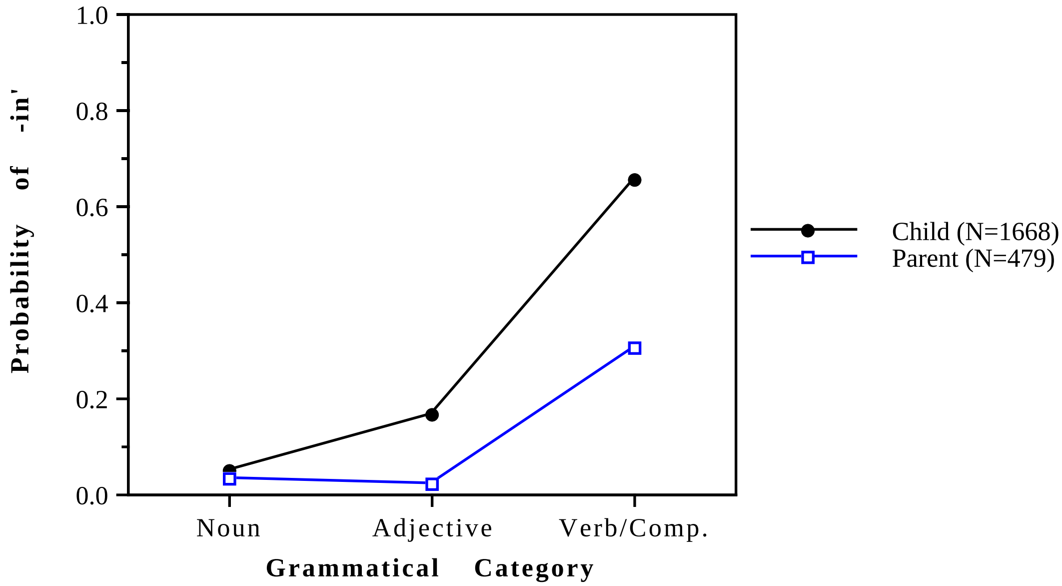
---

[29]As noted in Section 6.4.2, the children produced no true adjectives, but rather only gerunds used as adjectives, as in "swimming pool". These tokens make up the verbal adjective category.

[30]Instances of the words *something* and *nothing* were collected but not analyzed since they patterned very differently from other nouns. Most of the children had a high rate of deletion for these words; in fact, nine of them produced *something* and *nothing* categorically with [ɪn].

combined in Figure 9.1.

Figure 9.1: Effect of grammatical category on (ing) production



The individual results were quite similar to the group findings as shown in Table 9.1 and Figure 9.2. In Table 9.1, the percentages of the realization of (ing) as [ɪn] are presented. It is difficult to summarize the findings as to the range of [ɪn] production due to the low numbers of tokens in the noun, verbal adjective, and complement categories for the low MLU group, however, in general the results are consistent.[31] In the noun category, only one child used [ɪn] over 50% of the time (Danny, 67%), and the three nouns he produced were gerunds ("beepin'" as in "I hear a beepin'.") In the verbal adjective category, [ɪn] was also infrequently used, in that no child with more than two verbal adjective tokens produced [ɪn] more than 50% of the time with the exception of Zak, who will be discussed further below. On the other hand, the range of [ɪn] in the complement category was high -- from 67% to 100% for all of the

[31]See section 9.2 for a discussion of MLU, age, and (ing) production.

children except Cindy, who had only one complement token which she produced with [iN]. There were considerably more (ing) tokens in the verb category, and [ɪn] use was uniformly high with at range from 67% to 98%.

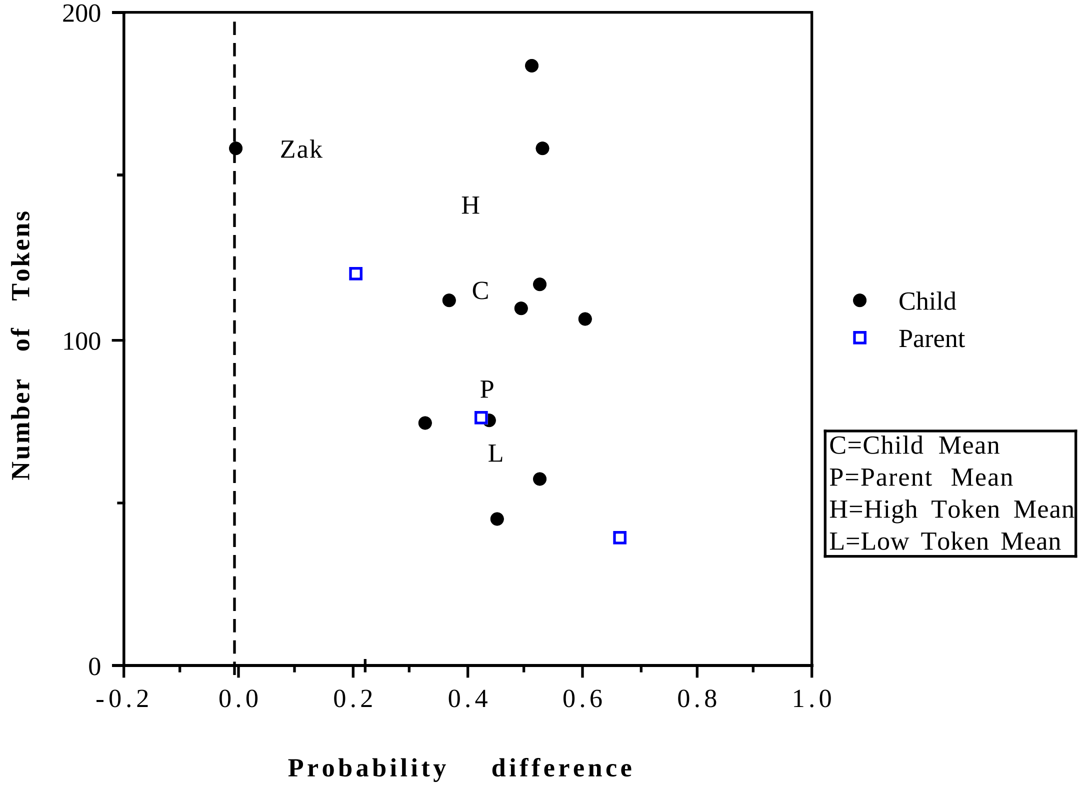
Table 9.1: Effect of grammatical form, age and MLU: Percentages of *in'* production.

NAME	AGE	MLU	NOUN N/%		V.ADJ. N/%		COMP. N/%		VERB N/%		TOTAL N/%
Low MLU											
Cindy	3-3	3.43	0	-	2	0	1	0	60	98	63/93
Diane	4-6	4.12	2	50	1	100	2	100	41	93	46/92
Mike	3-4	4.52	0	-	1	100	2	100	83	67	86/62
Marie	3-2	4.65	1	0	2	100	6	83	65	91	74/90
Micky	3-5	4.78	2	0	0	-	2	100	87	89	91/88
Evan	3-10	5.01	4	0	4	0	1	100	67	93	73/87
Rhea	3-11	5.17	5	0	1	100	2	100	49	82	59/76
Callie	3-11	5.30	0	-	4	50	1	100	70	91	75/90
Erin	3-3	5.33	0	-	0	-	1	100	43	93	44/93
High MLU											
Jeanie	3-9	5.38	3	33	4	50	3	67	164	94	174/89
Zak	4-11	5.53	2	50	7	86	7	71	135	78	151/78
Mira	4-3	5.59	6	17	2	100	0	-	94	76	112/68
Jenny	3-11	5.82	6	17	15	40	3	100	85	85	109/79
Kent	4-6	5.82	2	0	3	0	1	100	33	67	41/56
Danny	4-10	5.90	3	67	11	36	8	100	84	98	106/91
Shelly	4-9	6.29	6	0	5	40	8	75	139	90	158/86
Gia	3-11	6.31	7	29	2	0	1	100	105	89	115/83
Mean%			15.47		47.18		82.12		86.71		82.41
S.D.			22.26		42.61		32.98		9.61		48.30

In addition to the percentages listed above, individual variable rule analyses were performed on the data of the eleven children and three adults for whom there were enough tokens. The factors noun and verbal adjective were combined for the individual analyses because they patterned similarly for most of the children and were not significantly different from each other for the adults. As was stated in Chapter 7, the interpretation of the results in variable rule analysis relies not on the specific value of the probabilities themselves, but on the relationship between these probabilities. Therefore, individual results will again be shown by subtracting the probability of application (use of [ɪn]) for one grammatical category from the probability of application for another grammatical category. Specifically, in this case, the probability of [ɪn] in nouns/adjectives was subtracted from the probability of [ɪn] in verbs/complements. As was true in the analysis of (-t,d) deletion, a cluster in a given probability range would indicate a similar difference in effect of grammatical form in each of the speakers. The probability difference for each speaker is plotted against the total number of tokens per speaker in the vertical axis, since consistency of results might be expected to increase with number of tokens. As can be seen in Figure 9.2, however, this was not the case for the difference between the noun/verbal adjective and the complement/verb. The results were consistent regardless of total number of tokens. With only one exception, the difference between the effects of noun/adjective and complement/verb groups clustered between .2 and .7. For no subject was there a negative difference which would indicate the opposite effect of grammatical form from the one which would be predicted from previous research with adult speakers. The exception was a child (Zak) for whom the difference was .00. For the other factors affecting (ing) production, as well as for (-t,d) deletion, Zak's speech is very much like that of the other children, so there is no readily apparent explanation for this difference in the effect of grammatical form on (ing). Closer examination of the data reveals, however, that it is Zak's frequent use of [ɪn] in verbal adjectives that is the primary difference between him and the other children. The verbal adjective tokens consist, in this case, of six tokens of *cooking* as in "that cookin' stuff" all realized as [ɪn] and two of jumping in "jumping jacks" both realized as [iN]. The lack of diversity in the data suggest the possibility of a lexical explanation for the findings, or, at least, point out the need for larger amounts of data in interpreting them. The lexical explanation is also reinforced by the fact that one of Zak's nouns was also *cooking* which was produced with [ɪn].

The children were also divided into high and low token groups with 100 tokens as the dividing line between the groups. There was no significant difference between the mean probability differences of the two groups in the effect of grammatical form ($mean_{high\ token\ group}$=.432, $mean_{low\ token\ group}$=.435, t=.025, df=9.) This finding reinforces the above statement that the results are not affected by the number of tokens produced by each speaker. There was also no significant difference between the mean difference of the children and that of their parents ($mean_{children}$=.433, $mean_{parents}$=.431, t=.02, df=11.) Finally, there were no significant differences in variance for the groups (F=6.004, df=6,3 for high and low token groups; F=.056, df=10,2 for parents and children.)

Figure 9.2: Difference between the effects of noun/vrb. adj. and complement/verb in children and parents



The results of this analysis support the position that children are well on their way to learning the rule-governed variation of (ing) by the age of 3 to 4. As previously stated, the analysis of (-t,d) deletion, presented in the last chapter, gave clear evidence that the children were actually learning the deletion rule, not just copying words, with or without the final stop, from their parents. In that case, the children formed their own analysis of the semi-weak verbs (e.g. *felt, slept*) in which they deleted final stops from them as they did from monomorphemes. Their parents, on the other hand, were only as likely to delete (-t,d) from the semi-weak verbs as they were from the weak past tense verbs. This difference in (-t,d) deletion patterns indicates a difference in the analysis of the semi-weak form and, hence, the learning of rules, not individual forms, by the children. There is no such clear-cut case in which the children show a distinct break from their parents pattern with the variable (ing). Therefore, it is necessary to look at the individual tokens for indications as to whether the children are acquiring this alternation or copying individual forms from their parents. We might, however, infer that rule learning was taking place if a child used the same word as different parts of speech but produced it with [ɪn] in verbs or complements and with [iN] in nouns or gerunds attached to nouns. Unfortunately, there were not any uses of a single word across all four grammatical categories, but there were uses of a single word in different forms across two or three categories by two of the children. One of the activities done early on in the taping was a pretend cooking activity. This generated several used of the word *cooking*. Jenny used the word as a nominal activity label ("Cooking!") with an [iN] and as a complement ("We're all done cooking.") with an [ɪn]. Cindy used the word once as a noun ("The cooking.") with an [iN] referring to the activity itself, once as an gerund in the phrase "the cooking room" with an [iN] and twice with an [ɪn] as a progressive verb in the sentence "I'm cooking."[32] These examples work to support the evidence of the semi-weak verbs in the acquisition of

---

[32]Zak's tokens of *cooking*, discussed previously, are not consistent with the data reported on here. He used *cookin'* six times as a verbal adjective and once as a nominal. This difference between the children on the same word, again, suggests a lexical explanation for Zak's divergent probability difference illustrated in Figure 9.2.

the (-t,d) deletion rule that rule learning is what is occurring in these 3- and 4-year-old children.

The second question that must be addressed when talking about (ing) production in children is at what level the alternation between [ɪn] and [iN] is taking place. In other words, as Houston's (1985) work shows, (ing) variation is the contemporary result of an historical partial merger between *-ind* and *-and*, but one with no synchronic meaning attached to it. This, in addition, to the lack of evidence of a base form and a derived form suggest that, although (ing) variation acts very much like a rule, it is, rather, an alternation between two equivalent forms, [ɪn] and [iN]. There is nothing in the probabilities themselves that indicates whether (ing) is one morpheme or two for these children, and, with the present data, it is difficult to do more than suggest that [ɪn] and [iN] are allomorphs of the {ing} morpheme with different probabilities attached to each allomorph. One type of evidence for this interpretation would be the opposite of that indicating that the children are not learning individual forms, discussed above. In the present case, one would look for examples of uses of a single word in one grammatical category which vary in phonetic form between [iN] and [ɪn]. In fact, there are numerous examples of this in the data. For instance, Jenny uses *jumping* as a progressive verb with both [iN] and [ɪn]. Jeanie used *doing* in both forms as a present progressive verb. Dennis used *flying* as a verbal adjective in "a flying car" in both alveolar and velar forms. The children demonstrate by these examples that they know a single word can be produced in two ways. Additional data would be necessary to make any strong claims as to the level of the alternation, however, the suggestion is that [ɪn] and [iN] are two allomorphs of the {ing} morpheme with the verbal forms more likely to be associated with [ɪn] and the nominal forms more likely to be associated with [iN]. In addition, the fact that the grammatical form of the word influences the way in which it is produced by the children shows that they have made great progress in learning the variable rule-governed alternation of (ing).

9.2 Effect of age and MLU on (ing) production

In order to examine the data for possible effects of age or language level, the children were divided into groups by age and by mean length of utterance (MLU). Neither of these factors were chosen by the variable rule analysis program to have a significant effect on rule application (use of [ɪn]). Similarly, neither a comparison of the means of total rate of rule application nor one of mean use of [ɪn] in the verb category resulted in statistically significant findings ($t_{total\ means}$=1.297, ns, df=15; $t_{verb\ means}$=.834, ns, df=15.)[33] There was also no significant difference in dispersion of the data around the mean in either case (F_{total}=.771, ns, df=8,7; F_{verb}=.786, ns, df=8,7.) When the percentages of rule application for each child are examined, however, the differences in language level become more apparent. Table 9.2, below, shows the same

---

[33]As was also noted in the discussion of Figure 7.1 and can be seen in Tables 9.1 and 9.2, the standard deviations for these data are quite large. Again, the large variation in total numbers of tokens was influential. This was particularly true for the noun, verbal adjective, and complement categories, where data were sparse especially for the Low MLU group, and percentages ranged from 0% to 100%. Comparisons of these means would be meaningless and, therefore, was not attempted.

percentages of rule application as Table 9.1 but includes means and standard deviations for both the High and Low MLU groups.

Table 9.2: Effect of grammatical form, age and MLU: Percentages of *in'* production with means and standard deviations.

NAME	AGE	MLU	NOUN N/%	ADJ. N/%	COMP. N/%	VERB N/%	TOTAL N/%
Low MLU							
Cindy	3-3	3.43	0 -	2 0	1 0	60 98	63/93
Diane	4-6	4.12	2 50	1 100	2 100	41 93	46/92
Mike	3-4	4.52	0 -	1 100	2 100	83 67	86/62
Marie	3-2	4.65	1 0	2 100	6 83	65 91	74/90
Micky	3-5	4.78	2 0	0 -	2 100	87 89	91/88
Evan	3-10	5.01	4 0	4 0	1 100	67 93	73/87
Rhea	3-11	5.17	5 0	1 100	2 100	49 82	59/76
Callie	3-11	5.30	0 -	4 50	1 100	70 91	75/90
Erin	3-3	5.33	0 -	0 -	1 100	43 93	44/93
Mean%			10.00	64.29	87.00	88.56	85.67
S.D.			22.36	47.56	33.11	9.14	10.28
High MLU							
Jeanie	3-9	5.38	3 33	4 50	3 67	164 94	174/89
Zak	4-11	5.53	2 50	7 86	7 71	135 78	151/78
Mira	4-3	5.59	6 17	2 100	0 -	94 76	112/68
Jenny	3-11	5.82	6 17	15 40	3 100	85 85	109/79
Kent	4-6	5.82	2 0	3 0	1 100	33 67	41/56
Danny	4-10	5.90	3 67	11 36	8 100	84 98	106/91
Shelly	4-9	6.29	6 0	5 40	8 75	139 90	158/86
Gia	3-11	6.31	7 29	2 0	1 100	105 89	115/83
Mean%			26.63	44.00	87.57	84.63	78.75
S.D.			23.38	35.67	15.67	10.31	11.71
Total							
Mean%			15.47	47.18	82.12	86.71	82.41
S.D.			22.26	42.61	32.98	9.61	48.3

It is logical that there would be fewer tokens produced by the Low MLU group than by the High MLU group, and this turns out to be true in the present case. In addition, the Low MLU group results show far less variation in the noun, verbal adjective, and complement categories. Only three of the children in this group had anything but 0% or 100% rule application in any of these three categories. The tokens they do have, however, tend to show the predicted effects of the grammatical constraint. This finding is particularly striking in the noun column in Table 9.2. All of the Low MLU children but one have overall rates of use of alveolar (ing) between 56% and 93% across grammatical categories with all but three using [ɪn] over 75% of the time. In other words, they are all frequent users of the [ɪn] form. Yet, there is only one [ɪn] token for the Low MLU group in the noun category. In contrast, 16 of the 18 complement tokens are [ɪn] and there is a uniformly high percentage of [ɪn] in the verb column.

Table 9.3: Differences in numbers of tokens produced by High and Low MLU groups.

Group	Noun		V.Adj.*		Comp.		Verb		Total Tokens
	N	%	N	%	N	%	N	%	
Low MLU	14	2	15	2	18	3	565	92	611
High MLU	35	4	49	5	31	3	839	87	966

*x^2=15.41, p>.01 for verbal adjectives; other categories not significant.

The High MLU group produced far more of these nouns and a disproportionately high number of verbal adjectives as seen in Table

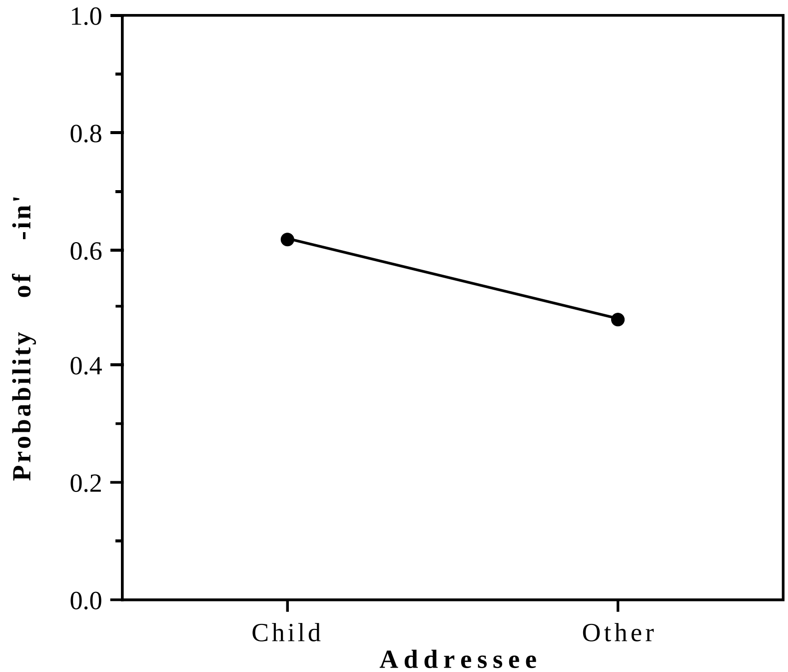
9.3. There is also more variation in rule application for these children. Nevertheless, Table 9.2 shows that they too demonstrate a low [ɪn] rate in the noun column. Danny is the only subject with an [ɪn] rate over 50%, and the two nouns he produced with [ɪn] were actually gerunds, as was previously stated. Therefore, the difference between the two groups of children is not in whether or not they demonstrate the grammatical constraints on (ing) variation, but in whether they use words in the relevant grammatical categories at all. It appears that when the (ing) words are used by the children, they are used in the form predicted by studies of adults. In other words, the youngest children use a high rate of [ɪn] in verbs, their most frequent (ing) words, but, when they begin using (ing) nouns and, to a lesser extent, verbal adjectives, they appear to use them with the "adult-like" [iN] form, not their own far more common [ɪn] form.

9.3 Effect of style on (ing) production

Unlike the results of the analysis of (-t,d) deletion in which the effect of addressee was not statistically significant, in the case of (ing) production, there was a significant effect of addressee on the form of (ing) used. As previously mentioned, during the interviews, the children talked to the researcher (or very rarely another adult), another child, or, at times, a puppet or other inanimate character. Although there was no difference in use of [ɪn] noted when the child addressed an adult versus an inanimate object, the children were significantly more likely to use the [ɪn] form with another child than with either of the other two addressees, as shown in Figure 9.3

(p<.01). The adults' speech was not analyzed for addressee since, in almost all cases, they spoke only to the researcher.

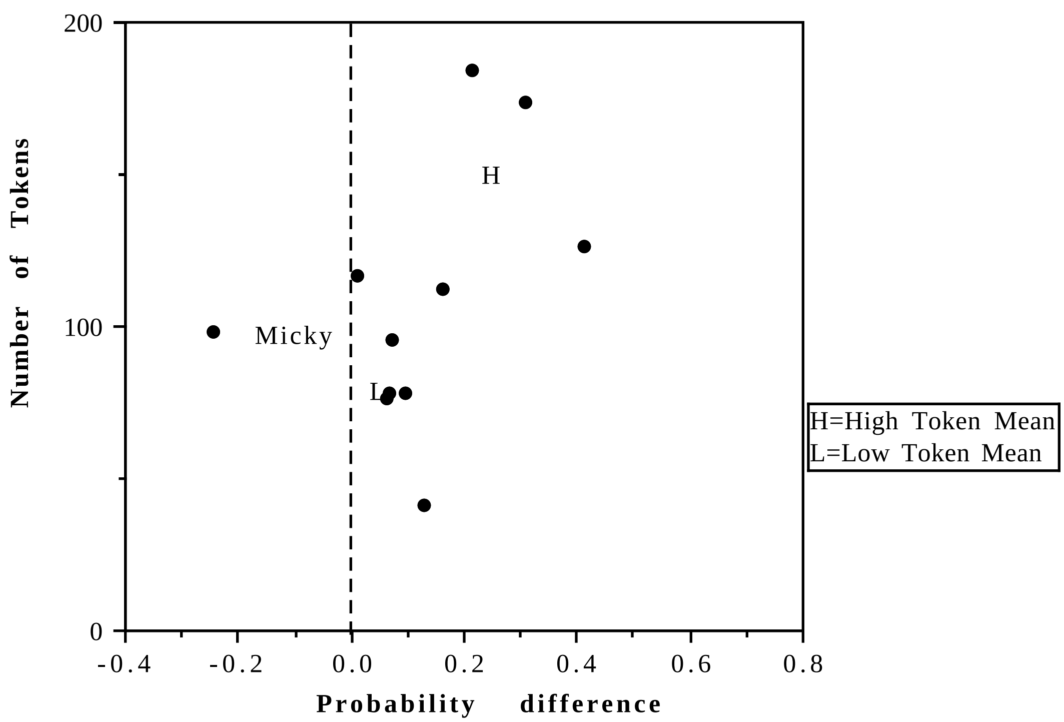
Figure 9.3: Effect of addressee on (ing) production by children (N=1688).



As shown in Figure 9.4, the results of the individual analyses again support those of the group analysis. In this case the probability of rule application when the addressee was an adult or an inanimate character was subtracted from that in which the addressee was a child. With only one exception, the children did show the relationship between them which was predicted from the group results. The exception was from the speech of one of the younger children (Micky), age 3 years 5 months, with a mid range number of tokens (N=98), for whom the difference between child and other addressee was -.246. This difference in addressee effect was opposite from the one predicted given the group results but was not statistically significant.

In the case of addressee, number of tokens was slightly but significantly correlated with the mean differences in effect of addressee ($mean_{high\ token\ group}$=.233, $mean_{low\ token\ group}$=.030, t=2.204, df=9, p<.05.) The significance of this finding held whether Micky's data were included or not. In other words, the children with a larger number of tokens demonstrated a greater effect of addressee on (ing) production than those with a smaller number of tokens. There was no significant difference, however, in the dispersion of the data about the means (F=1.211, df=4,5.)

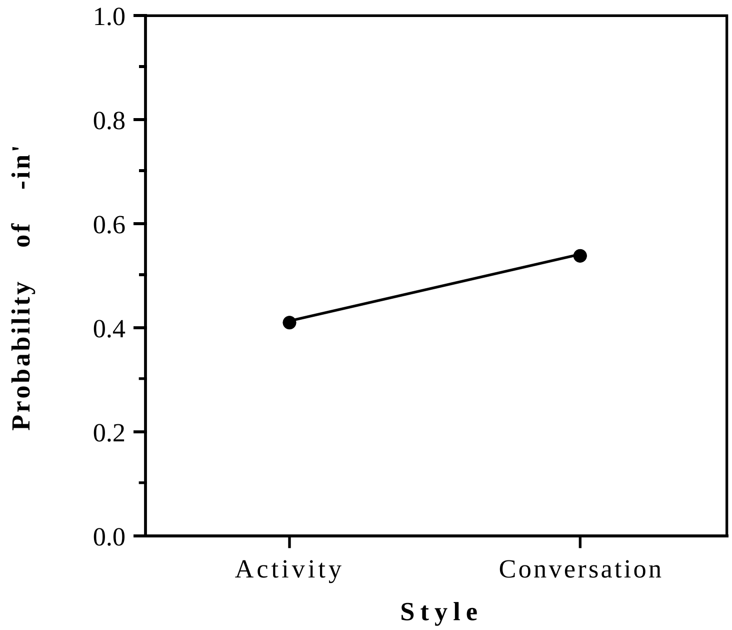
Figure 9.4: Differences between the effects of addressee on children. p(child addressee) - p(adult and "other" addressee)



The results of this analysis reveal that the children appear to have begun to acquire extralinguistic constraints on (ing) production based on addressee. It is impossible to interpret the results when the addressee is an inanimate object in light of adult results since, as one grows older, one presumably talks to fewer inanimate objects, at least during an interview session. The children do appear, however, to be making a distinction based on addressee similar to that found by Wald and Shopen (1985) in Los Angeles. They noted that speakers used more [ɪn] when talking to friends and family than to other addressees. The children in the current study also used more [ɪn] when talking to their classmates than when talking to the interviewer.

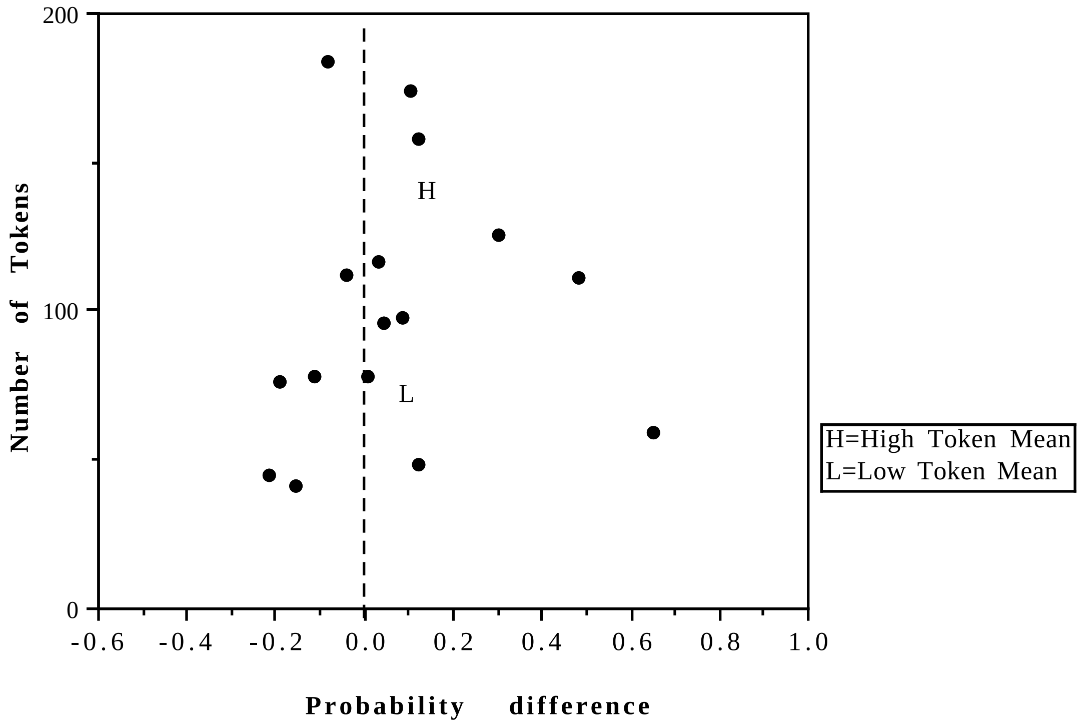
The results of this analysis of (ing) production were also examined based on the style of interaction. While Fischer (1958) found that the topic of conversation had an impact on the rate of use of the [ɪn] form for one 10-year-old boy, in the present study the styles were defined according to different activities within the interview. The styles examined were role playing in which a child pretended to be another person or character while playing with puppets or Sesame Street characters, pretend book reading, game playing (involving picture cards), and general conversation either with the researcher, another child, or an inanimate toy. The children did not vary their production significantly in response to any of the more structured activities or role playing, but all of these activities correlated with fewer instances of [ɪn] than general conversation as shown in Figure 9.5. The results were significant at the $p<.01$ level. The adult data was not examined for style since few narratives were told, and style demarcations comparable to those made for the children would not have been possible or appropriate for the adults.

Figure 9.5: Effect of style on (ing) production (N=1688).



The fit between individual and group results was not as clear-cut in the style analysis as in the analyses in the other categories. Although, as in the other factor groups, the differences appear to cluster around a range of probabilities, six of these differences are below zero, the opposite of the effect indicated by the group results. Thirteen of the sixteen individual analyses, including all of the ones resulting in a negative difference, were not statistically significant. Number of tokens was not significantly correlated with the size of the effect nor the dispersion of the differences around the group means ($mean_{high\ token\ group}$=.131, $mean_{low\ token\ group}$=.054, t=.619, df=13; F=.548, df=6,7.)

Figure 9.6: Difference between the effects of style on children. p(conversation) - p(activity)



Based on these findings, it must be concluded that the individual analyses of the style factor group are not as close a fit to the group analysis as was the case for the other factor groups. Therefore, one can infer that the style constraint has not been learned as well by the children as the grammatical and addressee constraints. As previously noted, however, the tokens were coded for style based on the activities in which the children participated. As was discussed in the analysis of (-t,d) deletion, it is possible that this breakdown of style factors did not capture differences in the effects of style that the children may have mastered, or that this situation was not one in which they showed consistent style effects.

These results, in combination with the lack of significant findings correlating style and (-t,d) deletion, suggest that style variation is acquired later than register. Although register was not formally examined in this study, the children were often noted to show signs of register change during the play activities. The Sesame Street play activity especially encouraged this type of role switch, which was signaled by a raising or lowering of vocal pitch and often increased volume depending on the character portrayed. As discussed in Chapter 3, it is not surprising that children of this age are engaging in changes in register during their play. There are frequent findings that preschool children can use the "baby talk" register, as well as portray other persons or roles which are familiar to them (e.g. parent, teacher, doctor, etc.) Style, however, requires a more subtle and systematic variation, in this case, in (ing) production or (-t,d) deletion. It appears from the present study that children are at the very early stages of its mastery.

In contrast to the findings on style, however, the acquisition of the constraints of grammatical form for (ing) production and (-t,d) deletion and phonological form for (-t,d) deletion are well on the way to completion. Together, these findings suggest that it is the grammatical and phonological constraints which are first learned by children as they acquire these variable rules and the stylistic constraints which come later.

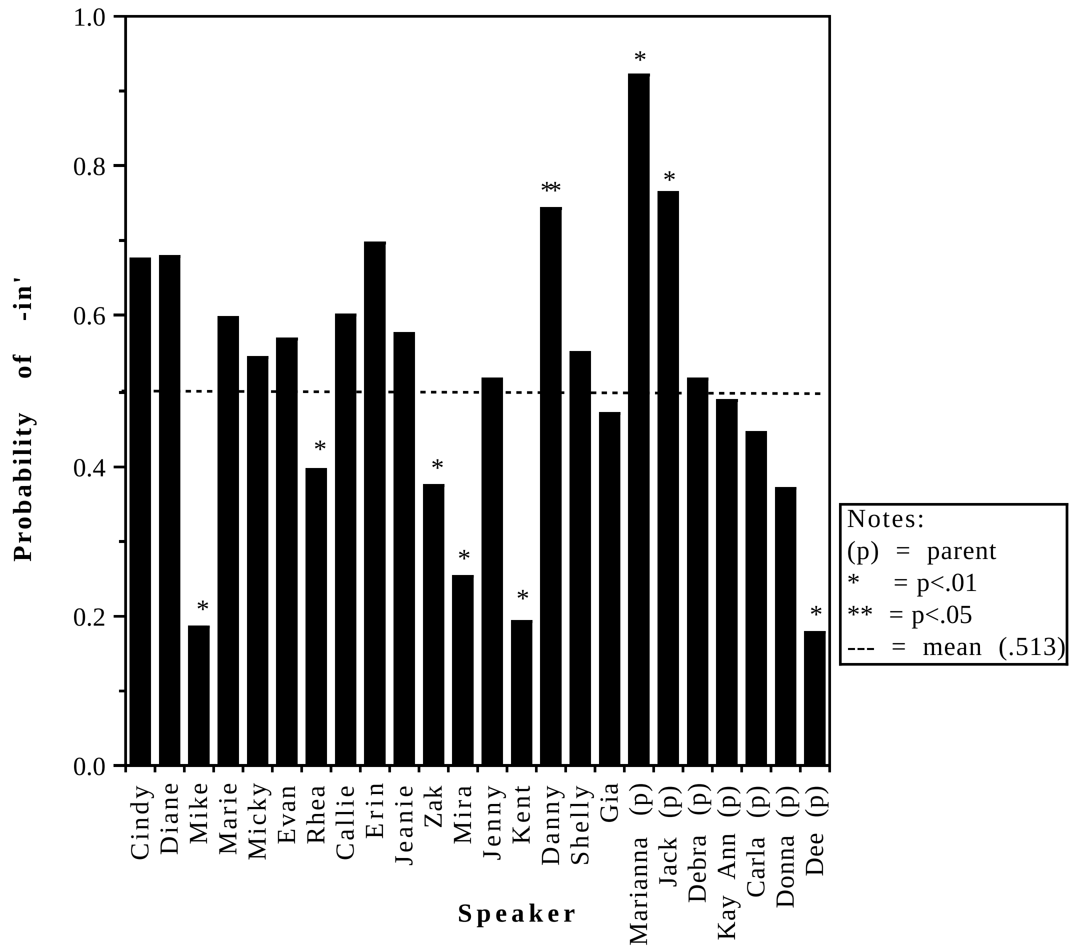
9.4 Effect of gender differences on (ing) production

Although the girls in this study were slightly more likely than the boys to use the [ɪn] form of (ing), these results were not statistically significant. As with (-t,d) deletion results, in which the girls were significantly more likely to delete (-t,d) than the boys, these results are interesting because they differ from previously reported findings that show girls to use fewer [ɪn] forms than boys (Fischer, 1958) and women to use fewer [ɪn] forms than men (Shuy, Wolfram, and Riley, 1968; Trudgill, 1974; Wald and Shopen, 1985; and Bradley and Bradley, 1979). Again, it appears that the girls have not as yet learned to be linguistically conservative.

9.5 Effect of individual differences on (ing) production

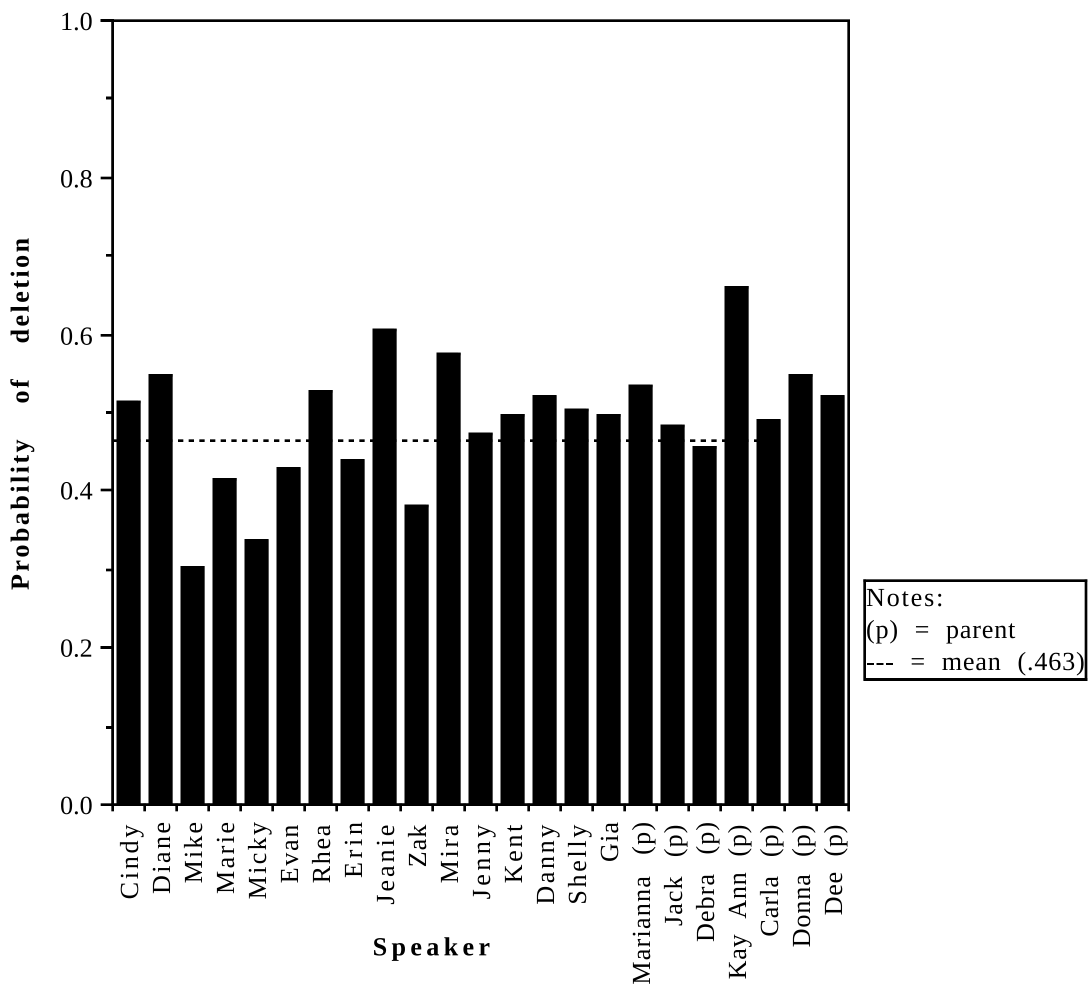
Although individual adults and children responded similarly to the various constraints, particularly those of grammatical form and audience, the overall rate of rule application differed among subjects. This was true of both children and parents as can be seen in Figure 9.7. Danny had a rate of [ɪn] use that was significantly higher than the other children's rate ($p<.05$). On the other hand, Rhea, Zak, Mira, Kent and Mike had rates of [ɪn] use that were significantly lower than the group's as a whole ($p<.01$). In the case of the parents, Jack (Gia's father) and Marianna (Callie's mother) had higher rates of [ɪn] use than the others, and Dee (Gia's mother) had a lower rate of [ɪn] use ($p<.01$).

Figure 9.7: Individual differences in (ing) production. Children are in order of MLU. (N=2147)



As was discussed in Chapter 7, individual differences in (-t,d) deletion were not significant. This finding is further illustrated below in Figure 9.8, where it can be seen that the individual probabilities of deletion hover closer to the mean deletion level than do those of production of [ɪn] (Figure 9.7). Whereas (-t,d) deletion is a variable rule which operates on word-final consonant clusters, Houston (1985) found that the different forms of (ing) are the historical residue of a fifteenth century partial merger. (See Chapter 5 for further discussion of her findings.) The present finding indicates that the influence of the variable rule for (-t,d) deletion is more consistent across individual speakers of both generations than that of the rule-governed alternation between two historically related allomorphs of {ing}.

Figure 9.8: Individual differences in (-t,d) deletion. (N=2444)



9.6 Summary

The results of the analysis of (ing) production in 3- and 4-year-old children showed a great similarity to previously documented adult speech and to that of the adults in the present study. In other words, as was the case with (-t,d) deletion, the children are well on their way to learning the variation in the production of alveolar [ɪn] and the grammatical constraint on that alternation. In addition, they are also beginning to acquire the extralinguistic constraints on (ing) production. These findings suggest that it is the grammatical and phonological constraints which are the first to be acquired. They are then followed by the extralinguistic or social constraints.

A review of specific findings on (ing) production includes the following:

1. The children's closest replication of the adult (ing) production pattern was in the area of grammatical form. The children were most likely to use the [ɪn] form in verbs and complements, less likely in gerunds acting as verbal adjectives, and least likely in nouns. In addition, the fact that there were instances in which the children used a word in different forms across two or three grammatical categories supported the findings for (-t,d) deletion that the children are learning rules/alternations, not copying individual forms when they produce the variants of (ing). More data is needed to make possible a stronger argument; however, there are also several examples of the children varying their production of (ing) and using both [iN] and [ɪn] for the same word in the same category. This suggests that the children are not learning these these different forms of (ing) as two separate morphemes, but, again, are learning a rule-like alternation governing the variation in two allomorphs of (ing).

2. There was no significant difference in rule application based on age or mean length of utterance (MLU), but the Low MLU children were less likely to use (ing) nouns and, particularly, verbal adjectives in their interviews than children in the High MLU group, and they were less likely to show variation in the noun, verbal adjective and complement categories. When they did use these forms, however, the tokens showed the effects of grammatical constraint which would be predicted given the findings for the High MLU group and the adults. This finding suggests that the children are learning the constraints on (ing) production as they learn the (ing) forms themselves and are in concert with similar findings on (-t,d) deletion as discussed in Chapter 7.
3. The children were more likely to use the [ɪn] form when addressing children than when addressing adults or inanimate characters.
4. Although the children were more likely to use [ɪn] in conversation than with any of the other activities provided during the taping sessions, these findings were less well supported by individual analyses than other results in (ing) production and (-t,d) deletion.
5. There were no significant sex differences with regard to (ing) production.
6. There were individual differences in the overall rate of rule application for both adults and children even though

the speakers responded similarly to the various constraints on (ing) production.

7. The findings are summarized in Table 9.4.

Table 9.4: Summary of results on (ing) production.

Constraint	Predicted Result	Actual Result
Grammatical form	verb>adj.>noun[34]	verb>ver. adj.>noun
Addressee	friend/family>other	child>other
Style	less formal>more formal	conversation>activity
Sex	male>female	not significant
Individual differences	n/a	some differences

[34]The symbol ">" is to be read "favors rule application more than."

CHAPTER 10: SUMMARY AND CONCLUSIONS

10.0 Goals of this study

There were several goals pursued during the present study. These include methodological goals as well as those to be addressed by the analysis. They are as follows:

1. To examine the acquisition of variation in children through a study of (-t,d) deletion and (ing) production in 3- and 4-year-olds.
2. To collect sufficient data to allow for individual variable rule analyses to supplement the group analyses and to develop research methods to gather the maximum amount of data containing (ing) and (-t,d) tokens as efficiently as possible in very young children.
3. To examine the data as it relates specifically to the following questions:
   a. When do children acquire variable rules, and more specifically, how does the time of acquisition of variable rules compare with that of the learning of related categorical rules?
   b. If 3- and 4-year-old children have acquired these variable rules and their constraints, can this acquisition be explained by natural, universal constraints on language production, or can it be the result of the acquisition of a socially learned dialect?

c. If the children demonstrate similar patterns of (-t,d) deletion and (ing) production to their parents, are they actually learning variable rules and their constraints, or are they matching the probabilities of the individual words they hear from the adults around them?

d. What are the segments which underlie the patterns of (-t,d) deletion and (ing) production for these children? In other words, are the (-t,d) segments really there in all of the grammatical forms in the grammar of these children? Can the learning of the two forms of (ing) be explained by the presence in the grammars of these children of two distinct (ing) morphemes, nominal {iN} and verbal {ɪn}?

10.1 Methodological considerations

As discussed in Chapter 6, several modifications of the sociolinguistic interview were necessary to adapt it for use with preschool children. The result was a play-interview session which contained the following components:

1. Adult/child conversation. This section of the session was initially led by my asking questions about their family and favorite activities, etc. but later evolved into a child-led dialogue about topics of the children's choice.
2. Role playing/toy manipulation. The children set up and played with a Sesame Street house which contained

furniture and characters. They then played with the toys often role playing with the various characters.

3. Child/child conversation. Toy telephones were used to facilitate this interaction which was generally led by the children.
4. Book "reading". The children were given both picture books and blank books and asked to "read" them. Verb tenses were encouraged by such questions as, "What happened then?" or "What's she doing now?"
5. Picture naming. Picture cards were drawn from a cloth bag and labeled by the children. When the bag was empty, the pictures were again labeled as the children tossed them back into the bag.

10.2 Summary of the discussion of theoretical considerations

10.2.1 Acquisition of variable and categorical rules

The dominant finding in both the analysis of (-t,d) deletion and (ing) production was that the children, at three and four years of age, have already made great progress in the acquisition of these two variable rules and the linguistic constraints that govern them. With the important exception of the semi-weak verbs and their effect on (-t,d) deletion, they are replicating the patterns of the adults with respect to both the grammatical constraints on both variables and the phonological and prosodic constraints on (-t,d) deletion. Among the implications of these findings are that the children are learning these variable rules at a very young age -- during the early language acquisition period. Are they, however, learning these rules concurrently with similar categorical rules? A methodological paradox makes a definitive answer to this question problematic. By the age of three, children have acquired the progressive verb form and are well on the way to acquiring the regular past tense forms, but collection of sufficient data for variable rule analysis from children under the age of three would be so time consuming as to be discouraging at best to the researcher. It might well take so long as to render the resulting data unusable due to the changes in the children's linguistic abilities during that period. Nevertheless, in the present study of 3- and 4-year-olds, there are clues to the answer.

In the (-t,d) deletion analysis, there were no significant differences in deletion among children based on age or mean length of utterance (MLU). However, the Low MLU group produces fewer of the more complex grammatical forms than the High MLU group. The fact that when they did produce these forms, they demonstrated conformity with the (-t,d) deletion rule, suggests that they acquire the (-t,d) deletion variable rule along with the related categorical rules.

The analysis of (-t,d) deletion also shows that the children have acquired both the grammatical and phonological constraints on (-t,d) deletion, even the dialectically dependent following pause constraint. They have also formed a different analysis of the semi-weak verb from their parents. At the same time, as would be expected given their ages and MLU’s, they have acquired a consistent weak past tense form even though they still evidence overregularization errors. In other words, the children have learned the basics of both a variable and a categorical rule, and, in both cases, they are in the process of learning the more advanced aspects of these rules.

Evidence from the analysis of (ing) production can also be brought to bear on this issue. At the time of the interviews, the children had all acquired the progressive *-ing* verb form. Although there were rare instances of omitted auxiliary verbs, there were no examples of a present tense verb form or verb root used in place of an obligatory progressive form. On the other hand, the three to four year age period, for these children, appeared to be the time when they began to use *-ing* forms as nouns, verbal adjectives, and complements. As with (-t,d) deletion, examination of the data shows that as they acquired these forms, they also acquired the variable rule that resulted in use of the form of (ing) that is predicted given results of adult (ing) research.

The findings in the current study suggest that the children are learning variable rules and the linguistic constraints on these rules at the same time as they are learning related categorical rules. While the preschool years have often been cited as the critical period of language development, discussions of this critical period have been limited to the acquisition of categorical rules. This study demonstrates, however, that the preschool years are also the critical period for the acquisition of variation and further suggests that the learning of variable rules and of categorical rules go hand-in-hand.

10.2.2 Other explanations of the current findings

There are several explanations other than the acquisition of variable rules which can be given for the findings presented in this study, and these have been discussed in detail in Chapters 7 and 9. The first is that natural and universal tendencies to preserve meaning and to favor ease of articulation work together to account for (-t,d) deletion. This explanation, however, cannot account for the effect of following pause on (-t,d) deletion. Guy (1980) points out that the following pause effect differs according to the dialect of the speaker and cannot be explained by ease of articulation or a tendency to preserve meaning, both of which are irrelevant in this case. The findings in the present study support this argument in that these children demonstrate that they are truly Philadelphia speakers by their acquisition of the following pause constraint. The fact that the children replicate the Philadelphia dialect pattern with respect to the effect of following segment suggests that they are learning rules grounded in a socially transmitted dialect rather than applying a universal constraint. At the same time, the fact that these children have their own deletion pattern as regards the semi-weak verb form shows that they are forming their own analysis of these forms, independent of their care-takers, not participating in a universal tendency to preserve meaning.

It might also be suggested, in discussing the findings on (ing) production, that the children are learning the [iN] and [ɪn] forms of (ing) as two separate morphemes. Although this possibility cannot be eliminated definitively, there is some preliminary evidence to suggest that the alternation is at the allomorph level since the children are able to use the same word with different (ing) forms in different grammatical categories, as discussed more fully in Chapter 9.

Finally, an explanation which might be brought forward as relevant to both (-t,d) deletion and (ing) production is that the children are not learning rules but simply matching probabilities attached to individual words spoken by their parents and other caretakers. Their treatment of the semi-weak verb speaks strongly against that possibility. These children demonstrate a radically different analysis of this verb form which could not possibly be acquired by imitating the individual forms heard from their parents. The most likely explanation for these findings is that the children are, in fact, learning variable rules and the linguistic constraints on those rules. Evidence presented in Chapter 9 that children use different forms of the same (ing) word across different grammatical categories gives further support to this argument. The history of the language acquisition literature shows that children learn rules and construct abstractions. The present study supports this body of work and argues strongly against a word-by-word learning approach. In fact, the present results show that children replicate their parents' linguistic pattern only in so far as they have the same theoretical base.

10.2.3 Extra-linguistic constraints on the variables

While the preschool period seems to be the critical one for learning the linguistic constraints on variable rules, the picture is not so clear for the extra-linguistic or social constraints. Stylistic conditioning based on addressee and activity were significant only for (ing) production. Even then, the effect of the activity on the form of (ing) used was not consistent for all of the children, suggesting a weaker finding than those for the linguistic constraints. Sex differences were significant only in the (-t,d) deletion analysis and ran directly counter to those predicted from previous adult studies. All of these findings suggest that extra-linguistic constraints on variable rules are acquired after linguistic constraints. Such findings do not seem surprising if one assumes that social constraints are learned by interacting with a variety of people, in a variety of situations, speaking on a variety of topics. The opportunities for these types of interactions would naturally increase as one grew older. Linguistic constraints, on the other hand, can be learned from as few as one member of a speech community. Advancing age would increase one's chance to practice this knowledge but would not increase the chances of learning it in the first place.

10.3 Conclusions

As is the case for categorical rules, the preschool period appears to be the critical one for learning the foundations of the variable rules of (-t,d) deletion and (ing) production. Some of the constraints on rules are refined in later years, even up through adulthood, as pointed out by Guy and Boyd (1990).[35] By the age of three and four, however, the many of these internal constraints have already been acquired, including the dialect specific following pause constraint on (-t,d) deletion, firmly establishing these children as

[35]Marcus et al. (1990), using data from Stemberger (1989) who collected a corpus of 7500 adult speech errors and assumptions about the proportion of irregular verbs used in casual conversation, came up with a crude estimate of an overregularization rate of .00004 overregularizations in 1000 sentences of casual speech in adults. This suggests that the learning of the categorical past tense rule also extends far beyond the early language learning period.

members of the Philadelphia speech community. The acquisition of social constraints on variation has its beginnings in early childhood, but the bulk of this learning appears to take place after the age of four.

Children at this critical stage of language acquisition have been under the researchers' microscopes for many years. However, the search has been for linguistic universals, with variation being largely ignored. The fact that at least some variable rules and their constraints are also learned during this preschool period demonstrates the importance of including variation in a complete picture of child language acquisition. Labov, Weinreich, and Herzog (1968) state that variation and its mastery is a part of "universal linguistic competence." It is similarly important to include variation in a model of the acquisition of this competence.

The results of this and other studies discussed in previous chapters suggest that study of both the similarities and the differences between child and adult language can do much to inform both the fields of sociolinguistics and language acquisition. Just as the inclusion of variation is necessary to form a complete picture of child language, so is the inclusion of children in variational research important in formulating a complete overall picture of variation in the speech community. It has always been a tenet of sociolinguistics that the language of the speech community is the source of linguistic theory, and it seems a natural extension to include in this practice the youngest members of that community.

APPENDIX A

Rules for calculating mean length of utterance (Brown, 1973: 54)

1. Start with the second page of the transcription unless that page involves a recitation of some kind. In this latter case start with the first recitation-free stretch. Count the first 100 utterances satisfying the following rules.

2. Only fully transcribed utterances are used; none with blanks. Portions of utterances, entered in parentheses to indicate doubtful transcription, are used.

3. Include all exact utterance repetitions. Stuttering is marked as repeated efforts at a single word; count the word once in the most complete form produced. In the few cases where a word is produced for emphasis or the like (*no, no, no*) count each occurrence.

4. Do not count such fillers as *mm* or *oh*, but do count *no, yeah* and *hi*.

5. All compound words (two or more free morphemes), proper names, and ritualized reduplications count as single words. Examples: *birthday, rackety-boom, choo-choo, quack-quack, night-night, pocketbook, see saw.* Justification is that no evidence that the constituent morphemes function as such for these children.

6. Count as one morpheme all irregular pasts of the verb (*got, did, went, saw*). Justification is that there is no evidence that the child relates these to present forms.

7. Count as one morpheme all diminutives (*doggie, mommie*) because these children at least do not seem to use the suffix productively. Diminutives are the standard forms used by the child.

8. Count as separate morphemes all auxiliaries (*is, have, will, can must, would*). Also all catenatives: *gonna, wanna, hafta*. These latter counted as single morphemes rather than as *going to* or *want to* because evidence is that they function so for the children. Count as separate morphemes all inflections, for

example, possessive {s}, plural {s}, third person singular {s}, regular past {d}, progressive {N}.

APPENDIX B

Sample variable rule analysis runs:

1. (-t,d) deletion, all subjects. Parents and children are combined with the factor groups of following segment, grammatical form, and syllable stress. Other significant groups are preceding segment and gender.

2. (-t,d) deletion, children only. Significant factor groups are following segment, grammatical form, syllable stress, preceding segment, and gender.

3. (ing) production, all subjects, includes addressee. Parents and children are combined with the grammatical form factor group. Other significant factor groups are addressee and individual differences.

4. (ing) production, children only, includes addressee. Significant factor groups are grammatical form, addressee, and individual differences.

5. (ing) production, all subjects, includes style. Parents and children are combined with the grammatical form factor group. Other significant groups are style and individual differences.

6. (ing) production, children only, includes style. Significant factor groups are grammatical form, style, and individual differences.

1. **(-t,d) deletion, all subjects.** Parents and children are combined with the factor groups of following segment, grammatical form, and syllable stress. Other significant groups are preceding segment and gender.

Number of cells: 114
Application value(s): 1
Total no. of factors: 24

Factor Group 1: Following Segment

Key to Factors (Column 1)

h = following pause (children)
f = following liquid and glide (children)
e = following obstruent (children)
g = following vowel (children)
c = following vowel (parents)
a = following obstruent (parents)
d = following pause (parents)
b = following liquid and glide (parents)

Group		Apps	Non-apps	Total	%
1					
h	N	207	615	822	34
	%	25	75		
f	N	151	186	337	14
	%	45	55		
e	N	205	134	339	14
	%	60	40		
g	N	83	260	343	14
	%	24	76		
c	N	12	95	107	4
	%	11	89		
a	N	115	63	178	7
	%	65	35		
d	N	41	122	163	7
	%	25	75		
b	N	64	91	155	6
	%	41	59		
Total	N	878	1566	2444	
	%	36	64		

Factor Group 2: Grammatical form

Key to Factors (Column 1)

o = monomorphemes (children)
n = contractions (children)
p = weak past tense verbs and participles (children)
q = semi-weak verbs (children)
k = monomorphemes (parents)
l = weak past tense verbs and participles (parents)
j = contractions (parents)
m = semi-weak verbs (parents)

Group		Apps	Non-Apps	Total	%
2					
o	N	305	507	812	33
	%	38	62		
n	N	237	192	429	18
	%	55	45		
p	N	74	463	537	22
	%	14	86		
q	N	30	33	63	3
	%	48	52		
k	N	121	143	264	11
	%	46	54		
l	N	45	133	178	7
	%	25	75		
j	N	58	71	129	5
	%	45	55		
m	N	8	24	32	1
	%	25	75		
Total	N	878	1566	2444	
	%	36	64		

Factor Group 3: Syllable Stress

Key to Factors (Column 1)

u = stressed syllables (children)
v = unstressed syllables (children)
s = stressed syllables (parents)
t = unstressed syllables (parents)

Group		Apps	Non-apps	Total	%
3					
u	N	362	976	1338	55
	%	27	73		
v	N	284	219	503	21
	%	56	44		
s	N	137	288	425	17
	%	32	68		
t	N	95	83	178	7
	%	53	47		
Total	N	878	1566	2444	
	%	36	64		

Factor Group 4: Preceding Consonant (children only)

Key to Factors (Column 1)

c = preceding consonant and syllable boundary
s = other

Group		Apps	Non-apps	Total	%
4					
c	N	143	23	166	7
	%	86	14		
s	N	735	1543	2278	93
	%	32	68		
Total	N	878	1566	2444	
	%	36	64		

Factor Group 5: Gender Differences (children only)

Key to Factors (Column 1)

g = girls
b = boys

Group		Apps	Non-apps	Total	%
5					
g	N	670	1063	1733	71
	%	39	61		
b	N	208	503	711	29
	%	29	71		
Total	N	878	1566	2444	
	%	36	64		
TOTAL	N	878	1566	2444	
	%	36	64		

Group	Factor	Weight	App/Total	Input&Weight
1:	h	0.373	0.25	0.23
	f	0.620	0.45	0.45
	e	0.722	0.60	0.57
	g	0.505	0.24	0.34
	c	0.171	0.11	0.09
	a	0.773	0.65	0.63
	d	0.305	0.25	0.18
	b	0.529	0.41	0.36
2:	o	0.629	0.38	0.46
	n	0.513	0.55	0.35
	p	0.278	0.14	0.16
	q	0.636	0.48	0.47
	k	0.670	0.46	0.51
	l	0.500	0.25	0.33
	j	0.255	0.45	0.15
	m	0.326	0.25	0.20
3:	u	0.445	0.27	0.29
	v	0.639	0.56	0.47
	s	0.484	0.32	0.32
	t	0.549	0.53	0.38
4:	c	0.893	0.86	0.81
	s	0.461	0.32	0.30
5:	g	0.524	0.39	0.36
	b	0.442	0.29	0.28

Cell	Total	App'ns	Expected	Error
hqvsg	1	0	0.465	0.870
hqusg	3	1	0.848	0.038
hqusb	2	0	0.443	0.569
hpvsg	28	6	4.495	0.601

hpvsb	17	2	2.060	0.002
hpusg	91	10	7.254	1.130
hpusb	71	6	4.173	0.850
hovsg	105	51	48.076	0.328
hovsb	55	18	20.813	0.612
housg	257	64	71.117	0.985
housb	149	39	32.207	1.828
houcg	1	0	0.789	3.739
hnvsg	4	0	1.377	2.099
hnvsb	1	0	0.274	0.378
hnvcg	7	4	5.858	3.612
hnvcb	2	2	1.574	0.541
hnusg	15	2	2.881	0.334
hnusb	12	1	1.756	0.381
hnucg	1	1	0.699	0.430
gqvsg	3	2	1.799	0.056
gqvsb	4	3	2.076	0.855
gqusg	16	7	6.465	0.074
gqusb	8	5	2.626	3.193
gqucg	1	1	0.869	0.151
gpvsg	10	4	2.475	1.250
gpvsb	8	0	1.533	1.896
gpusg	111	5	14.390	7.040
gpusb	76	5	7.369	0.843
gpucg	1	0	0.593	1.456
govsg	14	10	8.291	0.864
govsb	2	2	1.023	1.910
gousg	45	17	17.857	0.068
gousb	21	5	6.755	0.672
gnvsb	1	1	0.394	1.537
gnvcg	4	3	3.593	0.960
gnusg	13	11	3.772	19.508
gnusb	5	2	1.138	0.845
fqvsg	1	0	0.705	2.394
fqvsb	1	0	0.633	1.726
fqusg	7	2	3.642	1.543
fqusb	4	0	1.755	3.127
fpvsg	4	1	1.379	0.159
fpvsb	2	2	0.550	5.276
fpvcg	1	1	0.837	0.195
fpusg	32	5	6.156	0.269
fpusb	30	5	4.396	0.097
fovsg	11	6	7.689	1.233
fovsb	4	2	2.504	0.272
fousg	42	25	21.534	1.145
fousb	20	9	8.626	0.028
foucg	1	1	0.911	0.097
fnvsg	27	19	15.950	1.425
fnvsb	9	4	4.589	0.154
fnvcg	18	17	16.809	0.033
fnvcb	15	13	13.657	0.353
fnusg	53	26	20.955	2.009
fnusb	55	13	17.620	1.782
eqvsg	3	1	2.375	3.821

equsg	4	3	2.530	0.237
equsb	5	5	2.769	4.030
epvsg	9	6	4.094	1.627
epvsb	4	2	1.503	0.264
epusg	29	8	7.957	0.000
epusb	13	6	2.784	4.727
eovsg	20	14	15.732	0.894
eovsb	15	11	10.899	0.003
eousg	35	21	21.892	0.097
eousb	15	10	8.194	0.878
envsg	30	21	20.884	0.002
envsb	9	3	5.606	3.211
envcg	39	38	37.333	0.279
envcb	15	15	14.125	0.929
enusg	56	27	28.520	0.165
enusb	35	11	14.978	1.847
enucg	3	3	2.731	0.296
dmssg	5	1	0.430	0.828
dmssb	3	0	0.190	0.203
dltsg	10	3	2.014	0.604
dlssg	32	4	5.211	0.336
dlssb	2	1	0.246	2.636
dktsg	50	19	16.958	0.372
dktsb	2	1	0.540	0.536
dkssg	46	12	13.043	0.117
djtcg	4	0	1.829	3.370
djssg	9	0	0.561	0.599
cltsg	3	0	0.318	0.356
clssg	61	4	5.117	0.266
clssb	3	0	0.186	0.198
cktsg	6	3	1.167	3.571
ckssg	27	2	4.240	1.404
ckssb	6	2	0.710	2.656
cjtcg	1	1	0.284	2.521
bmssg	12	1	2.328	0.939
bltsg	5	2	1.962	0.001
bltsb	1	1	0.318	2.148
blssg	24	5	7.979	1.666
blssb	2	0	0.528	0.718
bktsg	22	10	12.492	1.151
bkssg	30	16	15.099	0.108
bkssb	2	0	0.844	1.461
bkscg	1	1	0.908	0.101
bjtcg	24	19	16.399	1.303
bjssg	31	8	4.509	3.162
bjscg	1	1	0.625	0.601
amssg	12	6	5.059	0.303
altsg	4	4	2.647	2.045
alssg	29	21	17.439	1.824
alssb	2	0	1.042	2.175
aktsg	19	9	15.185	12.545
akssg	50	45	37.713	5.730
akssb	3	1	2.066	1.767
ajtsg	1	1	0.401	1.496

```
 ajtcg    26          22        22.549     0.101
 ajssg    32           6        10.886     3.324
Total Chi-square = 174.3744
 Chi-square/cell = 1.5296
Log likelihood =  -1294.069
```

2. **(-t,d) deletion, children only.** Significant factor groups are following segment, grammatical form, syllable stress, preceding segment, and gender.

Number of cells: 75
Application value(s): 1
Total no. of factors: 14

Factor Group 1: Following Segment

Key to Factors (Column 1)

q = following pause
l = following liquid and glide
s = following consonant
v = following vowel

Group		Apps	Non-apps	Total	%
1					
q	N	207	615	822	45
	%	25	75		
l	N	151	186	337	18
	%	45	55		
s	N	205	134	339	18
	%	60	40		
v	N	83	260	343	19
	%	24	76		
Total	N	646	1195	1841	
	%	35	65		

Factor Group 2: Grammatical Form

Key to Factors (Column 1)

m = monomorphemes
c = contractions
p = weak past tense verbs and participles
d = semi-weak verbs

Group		Apps	Non-apps	Total	%
2					
m	N	305	507	812	44
	%	38	62		

c	N	237	192	429	23
	%	55	45		
p	N	74	463	537	29
	%	14	86		
d	N	30	33	63	3
	%	48	52		
Total	N	646	1195	1841	
	%	35	65		

Factor Group 3: Syllable Stress

Key to Factors (Column 1)

s = stressed syllable
u = unstressed syllable

Group		Apps	Non-apps	Total	%
3					
s	N	362	976	1338	73
	%	27	73		
u	N	284	219	503	27
	%	56	44		
Total	N	646	1195	1841	
	%	35	65		

Factor Group 4: Preceding Segment

Key to Factors (Column 1)

c = preceding consonant and syllable boundary
s = other

Group		Apps	Non-apps	Total	%
4					
c	N	99	10	109	6
	%	91	9		
s	N	547	1185	1732	94
	%	32	68		

Total	N	646	1195	1841	
	%	35	65		

Factor Group 5: Gender

Key to Factors (Column 1)

g = girls
b = boys

Group		Apps	Non-apps	Total	%
5					
g	N	444	712	1156	63
	%	38	62		
b	N	202	483	685	37
	%	29	71		
Total	N	646	1195	1841	
	%	35	65		
TOTAL	N	646	1195	1841	
	%	35	65		

Group	Factor	Weight	App/Total	Input&Weight
1:	q	0.365	0.25	0.22
	l	0.611	0.45	0.43
	s	0.714	0.60	0.55
	v	0.497	0.24	0.32
2:	m	0.631	0.38	0.45
	c	0.521	0.55	0.35
	p	0.280	0.14	0.16
	d	0.639	0.48	0.46
3:	s	0.445	0.27	0.28
	u	0.642	0.56	0.47
4:	c	0.877	0.91	0.77
	s	0.469	0.32	0.30
5:	g	0.530	0.38	0.35
	b	0.450	0.29	0.28

Cell	Total	App'ns	Expected	Error
vpusg	10	4	2.496	1.208
vpusb	8	0	1.556	1.932

vpssg	111	5	14.361	7.009
vpssb	76	5	7.401	0.863
vpscg	1	0	0.544	1.193
vmusg	14	10	8.317	0.839
vmusb	2	2	1.030	1.882
vmssg	45	17	17.792	0.058
vmssb	21	5	6.760	0.676
vdusg	3	2	1.808	0.052
vdusb	4	3	2.096	0.820
vdssg	16	7	6.461	0.075
vdssb	8	5	2.637	3.158
vdscg	1	1	0.845	0.184
vcusb	1	1	0.403	1.480
vcucg	4	3	3.528	0.669
vcssg	13	11	3.818	19.132
vcssb	5	2	1.159	0.794
spusg	9	6	4.108	1.603
spusb	4	2	1.515	0.250
spssg	29	8	7.913	0.001
spssb	13	6	2.783	4.730
smusg	20	14	15.740	0.903
smusb	15	11	10.927	0.002
smssg	35	21	21.799	0.078
smssb	15	10	8.178	0.892
sdusg	3	1	2.379	3.858
sdssg	4	3	2.524	0.243
sdssb	5	5	2.770	4.026
scusg	30	21	21.043	0.000
scusb	9	3	5.674	3.409
scucg	39	38	37.037	0.498
scucb	15	15	13.979	1.095
scssg	56	27	28.681	0.202
scssb	35	11	15.138	1.994
scscg	3	3	2.682	0.356
qpusg	28	6	4.540	0.560
qpusb	17	2	2.094	0.005
qpssg	91	10	7.244	1.139
qpssb	71	6	4.195	0.826
qmusg	105	51	48.291	0.281
qmusb	55	18	21.012	0.699
qmssg	257	64	70.838	0.911
qmssb	149	39	32.250	1.803
qmscg	1	0	0.753	3.055
qdusg	1	0	0.469	0.882
qdssg	3	1	0.848	0.038
qdssb	2	0	0.445	0.572
qcusg	4	0	1.405	2.166
qcusb	1	0	0.282	0.393
qcucg	7	4	5.691	2.687
qcucb	2	2	1.519	0.634
qcssg	15	2	2.922	0.361
qcssb	12	1	1.793	0.412
qcscg	1	1	0.660	0.515
lpusg	4	1	1.384	0.163

lpusb	2	2	0.555	5.207
lpucg	1	1	0.809	0.235
lpssg	32	5	6.118	0.253
lpssb	30	5	4.394	0.098
lmusg	11	6	7.694	1.242
lmusb	4	2	2.513	0.282
lmssg	42	25	21.413	1.226
lmssb	20	9	8.605	0.032
lmscg	1	1	0.893	0.120
ldusg	1	0	0.707	2.411
ldusb	1	0	0.636	1.751
ldssg	7	2	3.631	1.521
ldssb	4	0	1.756	3.129
lcusg	27	19	16.113	1.283
lcusb	9	4	4.661	0.195
lcucg	18	17	16.603	0.122
lcucb	15	13	13.442	0.140
lcssg	53	26	21.097	1.893
lcssb	55	13	17.840	1.944

Total Chi-square = 107.3477
Chi-square/cell = 1.4313
Log likelihood = -978.550

3. **(ing) production, all subjects, includes addressee.** Parents and children are combined with the grammatical form factor group. Other significant factor groups are addressee and individual differences.

Number of cells: 129
Application value(s): 1
Total no. of factors: 35

Factor Group 1: Grammatical Form

Key to Factors (Column 1)

g = nouns (children)
e = verbs and complements (children)
f = verbal adjectives (children)
h = something/nothing (children)
b = verbal adjectives (parents)
a = verbs and complements (parents)
c = nouns (parents)
d = something/nothing (parents)

Group		Apps	Non-apps	Total	%
1					
g	N	9	40	49	2
	%	18	82		
e	N	1261	192	1453	68
	%	87	13		
f	N	29	32	61	3
	%	48	52		
h	N	76	29	105	5
	%	72	28		
b	N	4	33	37	2
	%	11	89		
a	N	186	166	352	16
	%	53	47		
c	N	7	53	60	3
	%	12	88		
d	N	15	15	30	1
	%	50	50		
Total	N	1587	560	2147	
	%	74	26		

Factor Group 2: Addressee (Children only)

Key to Factors (Column 1)

k = children
o = others

Group		Apps	Non-apps	Total	%
2					
o	N	1349	513	1862	87
	%	72	28		
k	N	238	47	285	13
	%	84	16		
Total	N	1587	560	2147	
	%	74	26		

Factor Group 3: Individual Differences

Key to Factors (Column 1)

j = Jenny
p = Cindy
d = Danny
k = Jeanie
s = Zak
l = Callie
g = Gia
v = Evan
h = Shelly
a = Diane
b = Mike
c = Micky
m = Marie
r = Erin
i = Mira
y = Kent
t = Rhea
e = Denise (mother of Gia)
f = Marianna (mother of Callie)
o = Kay Ann (grandmother of Jenny)
w = Carla (mother of Evan)
n = Donna (mother of Danny)
z = Jack (father of Gia)
q = Mary (mother of Jeanie)
u = Debra (mother of Shelly)

Group		Apps	Non-apps	Total	%
3					
j	N	99	27	126	6
	%	79	21		
p	N	62	5	67	3
	%	93	7		
d	N	101	10	111	5
	%	91	9		
k	N	163	21	184	9
	%	89	11		
s	N	123	35	158	7
	%	78	22		
l	N	70	8	78	4
	%	90	10		
g	N	97	20	117	5
	%	83	17		
v	N	66	10	76	4
	%	87	13		
h	N	149	25	174	8
	%	86	14		
a	N	44	4	48	2
	%	92	8		
b	N	59	37	96	4
	%	61	39		
c	N	86	12	98	5
	%	88	12		
m	N	70	8	78	4
	%	90	10		
r	N	42	3	45	2
	%	93	7		
i	N	76	36	112	5
	%	68	32		
y	N	23	18	41	2
	%	56	44		
t	N	45	14	59	3
	%	76	24		

e	N	6	33	39	2
	%	15	85		
f	N	32	7	39	2
	%	82	18		
o	N	17	27	44	2
	%	39	61		
w	N	32	47	79	4
	%	41	59		
n	N	43	82	125	6
	%	34	66		
z	N	24	14	38	2
	%	63	37		
q	N	11	12	23	1
	%	48	52		
u	N	47	45	92	4
	%	51	49		
Total	N	1587	560	2147	
	%	74	26		
TOTAL	N	1587	560	2147	
	%	74	26		

Group	Factor	Weight	App/Total	Input&Weight
1:	g	0.051	0.18	0.17
	e	0.653	0.87	0.88
	f	0.168	0.48	0.43
	h	0.448	0.72	0.75
	b	0.023	0.11	0.08
	a	0.304	0.53	0.62
	c	0.034	0.12	0.12
	d	0.287	0.50	0.60
2:	o	0.484	0.72	0.78
	k	0.604	0.84	0.85
3:	j	0.523	0.79	0.80
	p	0.696	0.93	0.90
	d	0.759	0.91	0.92
	k	0.597	0.89	0.85
	s	0.394	0.78	0.71
	l	0.621	0.90	0.86
	g	0.495	0.83	0.79

v	0.591	0.87	0.84
h	0.570	0.86	0.83
a	0.700	0.92	0.90
b	0.194	0.61	0.47
c	0.561	0.88	0.83
m	0.619	0.90	0.86
r	0.716	0.93	0.90
i	0.272	0.68	0.58
y	0.197	0.56	0.48
t	0.419	0.76	0.73
e	0.146	0.15	0.39
f	0.901	0.82	0.97
o	0.407	0.39	0.72
w	0.373	0.41	0.69
n	0.308	0.34	0.62
z	0.712	0.63	0.90
q	0.416	0.48	0.73
u	0.452	0.51	0.76

Cell	Total	App'ns	Expected	Error
hoy	2	0	0.823	1.397
hov	3	3	2.413	0.730
hot	1	1	0.672	0.488
hos	7	6	4.540	1.336
hop	4	3	3.466	0.469
hom	1	1	0.822	0.217
hol	3	3	2.469	0.645
hok	9	4	7.271	7.657
hoj	15	15	11.359	4.807
hoi	10	2	5.155	3.984
hoh	16	16	12.643	4.248
hog	1	0	0.735	2.780
hod	5	5	4.498	0.558
hoc	6	6	4.706	1.650
hob	9	0	3.659	6.166
hoa	3	3	2.606	0.453
hkt	1	1	0.769	0.300
hkm	3	3	2.648	0.399
hkk	1	0	0.872	6.840
hkj	2	2	1.671	0.394
hkg	1	1	0.819	0.221
hkc	1	1	0.855	0.169
hkb	1	0	0.527	1.115
goy	2	0	0.088	0.093
gov	3	0	0.642	0.816
got	5	0	0.598	0.679
gos	2	1	0.218	3.151
gor	1	1	0.322	2.104
gom	1	0	0.234	0.306
gok	2	0	0.436	0.557
goj	4	0	0.685	0.827
goi	6	1	0.395	0.992
goh	6	0	1.198	1.497
gog	6	2	0.933	1.444

god	3	2	1.118	1.110
goc	1	0	0.194	0.241
goa	1	0	0.305	0.439
gkv	1	0	0.307	0.443
gkk	1	1	0.312	2.207
gkj	2	1	0.503	0.655
gkg	1	0	0.231	0.300
gkc	1	0	0.282	0.392
foy	2	0	0.296	0.348
fov	1	0	0.506	1.022
fot	1	1	0.338	1.960
fos	7	6	2.204	9.544
fop	2	0	1.235	3.232
fom	2	2	1.070	1.740
fol	4	2	2.146	0.021
fok	4	2	2.046	0.002
foj	13	4	5.684	0.886
foi	2	2	0.419	7.551
foh	5	2	2.420	0.141
fog	1	0	0.409	0.692
fod	11	4	7.596	5.502
fob	1	1	0.146	5.862
foa	1	1	0.622	0.607
fky	1	0	0.221	0.283
fkj	2	2	1.117	1.582
fkg	1	0	0.530	1.126
eoy	29	19	17.945	0.163
eov	51	47	46.163	0.160
eot	43	34	35.537	0.383
eos	128	96	103.793	3.094
eor	36	33	33.962	0.481
eop	57	55	53.455	0.718
eom	62	56	56.714	0.106
eol	58	53	53.086	0.002
eok	150	139	136.070	0.679
eoj	73	60	64.151	2.216
eoi	59	44	42.006	0.329
eoh	125	110	112.183	0.414
eog	73	65	63.212	0.377
eod	92	90	87.786	1.219
eoc	79	71	70.640	0.017
eob	62	42	38.078	1.047
eoa	39	36	36.620	0.172
eky	5	4	3.627	0.140
ekv	17	16	15.971	0.001
ekt	8	8	7.085	1.033
eks	14	14	12.245	2.007
ekr	8	8	7.715	0.295
ekp	4	4	3.843	0.163
ekm	9	8	8.512	0.569
ekl	13	12	12.300	0.136
ekk	17	17	15.994	1.070
ekj	15	15	13.828	1.272
eki	35	27	28.029	0.190

ekh	22	21	20.556	0.146
ekg	33	29	30.132	0.490
ekc	10	8	9.322	2.766
ekb	23	16	16.593	0.076
eka	4	4	3.846	0.160
doz	3	2	2.330	0.210
dow	3	0	1.368	2.513
dou	6	3	3.229	0.035
doq	5	2	2.507	0.205
doo	7	4	3.443	0.177
don	5	3	1.928	0.971
doe	1	1	0.194	4.145
coz	3	0	0.700	0.913
cow	9	1	0.614	0.260
cou	7	0	0.647	0.713
coq	1	0	0.081	0.088
coo	8	0	0.624	0.677
con	20	3	1.039	3.901
cof	6	3	3.177	0.021
coe	6	0	0.124	0.126
boz	7	0	1.189	1.432
bow	7	1	0.328	1.441
bou	3	0	0.192	0.205
boq	1	0	0.056	0.059
boo	5	0	0.269	0.284
bon	5	2	0.178	19.358
bof	4	1	1.723	0.533
boe	5	0	0.070	0.071
aoz	25	22	19.777	1.196
aow	51	26	24.317	0.223
aou	72	41	40.242	0.032
aoq	16	9	8.357	0.103
aoo	21	10	10.770	0.113
aon	84	31	34.071	0.466
aof	28	27	26.135	0.430
aoe	27	5	5.612	0.084
akw	9	4	5.375	0.873
aku	4	3	2.693	0.107
ako	3	3	1.894	1.752
akn	11	4	5.787	1.165
akf	1	1	0.958	0.044

Total Chi-square = 174.3929
Chi-square/cell = 1.3519
Log likelihood = -909.051

4. **(ing) production, children only, includes addressee.** Significant factor groups are grammatical form, addressee, and individual differences.

Number of cells: 93
Application value(s): 1
Total no. of factors: 23

Factor Group 1: Grammatical Form

Key to Factors (Column 1)

n = nouns
v = verbs and complements
a = verbal adjectives
s = something/nothing

Group		Apps	Non-apps	Total	%
1					
n	N	9	40	49	3
	%	18	82		
v	N	1261	192	1453	87
	%	87	13		
a	N	29	32	61	4
	%	48	52		
s	N	76	29	105	6
	%	72	28		
Total	N	1375	293	1668	
	%	82	18		

Factor Group 2: Addressee

Key to Factors (Column 1)

o = others
c = children

Group		Apps	Non-apps	Total	%
2 (3)					
o	N	1152	259	1411	85
	%	82	18		
k	N	223	34	257	15

	%	87	13		
Total	N	1375	293	1668	
	%	82	18		

Factor Group 3: Individual Differences

Key to Factor Groups (See Run #3, Factor Group 3)

Group		Apps	Non-apps	Total	%
3 (5)					
j	N	99	27	126	8
	%	79	21		
p	N	62	5	67	4
	%	93	7		
d	N	101	10	111	7
	%	91	9		
k	N	163	21	184	11
	%	89	11		
s	N	123	35	158	9
	%	78	22		
l	N	70	8	78	5
	%	90	10		
g	N	97	20	117	7
	%	83	17		
v	N	66	10	76	5
	%	87	13		
h	N	149	25	174	10
	%	86	14		
a	N	44	4	48	3
	%	92	8		
b	N	59	37	96	6
	%	61	39		
c	N	86	12	98	6
	%	88	12		
m	N	70	8	78	5
	%	90	10		
r	N	42	3	45	3
	%	93	7		

i	N	76	36	112	7
	%	68	32		
y	N	23	18	41	2
	%	56	44		
t	N	45	14	59	4
	%	76	24		
Total	N	1375	293	1668	
	%	82	18		
TOTAL	N	1375	293	1668	
	%	82	18		

Group	Factor	Weight	App/Total	Input&Weight
1:	n	0.035	0.18	0.18
	v	0.559	0.87	0.89
	a	0.120	0.48	0.45
	s	0.355	0.72	0.77
2:	o	0.478	0.82	0.85
	k	0.618	0.87	0.91
3:	j	0.503	0.79	0.86
	p	0.680	0.93	0.93
	d	0.746	0.91	0.95
	k	0.578	0.89	0.89
	s	0.376	0.78	0.79
	l	0.602	0.90	0.90
	g	0.472	0.83	0.84
	v	0.571	0.87	0.89
	h	0.551	0.86	0.88
	a	0.684	0.92	0.93
	b	0.181	0.61	0.57
	c	0.542	0.88	0.88
	m	0.600	0.90	0.90
	r	0.700	0.93	0.93
	i	0.255	0.68	0.68
	y	0.185	0.56	0.58
	t	0.400	0.76	0.80

Cell	Total	App'ns	Expected	Error
voy	29	19	17.871	0.186
vov	51	47	46.100	0.183
vot	43	34	35.477	0.351
vos	128	96	103.684	2.997
vor	36	33	33.945	0.461
vop	57	55	53.441	0.728
vom	62	56	56.664	0.090
vol	58	53	53.043	0.000

vok	150	139	135.983	0.716
voj	73	60	64.045	2.083
voi	59	44	41.747	0.416
voh	125	110	112.097	0.380
vog	73	65	63.045	0.444
vod	92	90	87.770	1.232
voc	79	71	70.579	0.024
vob	62	42	37.819	1.185
voa	39	36	36.608	0.165
vky	5	4	3.695	0.097
vkv	17	16	16.033	0.001
vkt	8	8	7.141	0.963
vks	14	14	12.356	1.863
vkr	8	8	7.734	0.275
vkp	4	4	3.854	0.151
vkm	9	8	8.544	0.682
vkl	13	12	12.345	0.192
vkk	17	17	16.061	0.994
vkj	15	15	13.897	1.190
vki	35	27	28.352	0.339
vkh	22	21	20.651	0.096
vkg	33	29	30.286	0.664
vkc	10	8	9.366	3.141
vkb	23	16	16.877	0.171
vka	4	4	3.857	0.148
soy	2	0	0.820	1.390
sov	3	3	2.408	0.737
sot	1	1	0.671	0.490
sos	7	6	4.540	1.336
sop	4	3	3.467	0.471
som	1	1	0.821	0.218
sol	3	3	2.467	0.648
sok	9	4	7.269	7.643
soj	15	15	11.338	4.845
soi	10	2	5.116	3.885
soh	16	16	12.639	4.255
sog	1	0	0.733	2.742
sod	5	5	4.499	0.557
soc	6	6	4.704	1.654
sob	9	0	3.633	6.093
soa	3	3	2.607	0.453
skt	1	1	0.782	0.278
skm	3	3	2.670	0.370
skk	1	0	0.881	7.400
skj	2	2	1.690	0.367
skg	1	1	0.829	0.207
skc	1	1	0.865	0.156
skb	1	0	0.544	1.193
noy	2	0	0.087	0.092
nov	3	0	0.634	0.804
not	5	0	0.592	0.672
nos	2	1	0.217	3.177
nor	1	1	0.320	2.125
nom	1	0	0.232	0.303

nok	2	0	0.433	0.553
noj	4	0	0.677	0.815
noi	6	1	0.387	1.038
noh	6	0	1.190	1.485
nog	6	2	0.917	1.509
nod	3	2	1.115	1.119
noc	1	0	0.193	0.239
noa	1	0	0.304	0.436
nkv	1	0	0.321	0.472
nkk	1	1	0.328	2.053
nkj	2	1	0.528	0.572
nkg	1	0	0.241	0.318
nkc	1	0	0.296	0.421
aoy	2	0	0.295	0.346
aov	1	0	0.503	1.014
aot	1	1	0.337	1.968
aos	7	6	2.204	9.543
aop	2	0	1.236	3.237
aom	2	2	1.067	1.747
aol	4	2	2.142	0.020
aok	4	2	2.044	0.002
aoj	13	4	5.659	0.861
aoi	2	2	0.414	7.669
aoh	5	2	2.418	0.140
aog	1	0	0.406	0.682
aod	11	4	7.601	5.521
aob	1	1	0.144	5.933
aoa	1	1	0.623	0.606
aky	1	0	0.234	0.305
akj	2	2	1.152	1.472
akg	1	0	0.546	1.203

Total Chi-square = 130.4701
Chi-square/cell = 1.4029
Log likelihood = -641.709

5. **(ing) production, all subjects, includes style.** Parents and children are combined with the grammatical form factor group. Other significant groups are style and individual differences.

Number of cells: 134
Application value(s): 1
Total no. of factors: 35

Factor Group 1: Grammatical Form

Key to Factors (Column 1)

g = nouns (children)
e = verbs and complements (children)
f = verbal adjectives (children)
h = something/nothing (children)
b = verbal adjectives (parents)
a = verbs and complements (parents)
c = nouns (parents)
d = something/nothing (parents)

Group		Apps	Non-apps	Total	%
1					
g	N	9	40	49	2
	%	18	82		
e	N	1261	192	1453	68
	%	87	13		
f	N	29	32	61	3
	%	48	52		
h	N	76	29	105	5
	%	72	28		
b	N	4	33	37	2
	%	11	89		
a	N	186	166	352	16
	%	53	47		
c	N	7	53	60	3
	%	12	88		
d	N	15	15	30	1
	%	50	50		
Total	N	1587	560	2147	
	%	74	26		

Factor Group 2: Style (children only)

Key to Factors (Column 1)

o = conversation
n = activity

Group		Apps	Non-apps	Total	%
2					
o	N	1180	434	1614	75
	%	73	27		
n	N	407	126	533	25
	%	76	24		
Total	N	1587	560	2147	
	%	74	26		

Factor Group 3: Individual Differences

Key to Factors (See Run #3, Factor Group 3)

Group		Apps	Non-apps	Total	%
3					
j	N	99	27	126	6
	%	79	21		
p	N	62	5	67	3
	%	93	7		
d	N	101	10	111	5
	%	91	9		
k	N	163	21	184	9
	%	89	11		
s	N	123	35	158	7
	%	78	22		
l	N	70	8	78	4
	%	90	10		
g	N	97	20	117	5
	%	83	17		
v	N	66	10	76	4
	%	87	13		

h	N	149	25	174	8
	%	86	14		
a	N	44	4	48	2
	%	92	8		
b	N	59	37	96	4
	%	61	39		
c	N	86	12	98	5
	%	88	12		
m	N	70	8	78	4
	%	90	10		
r	N	42	3	45	2
	%	93	7		
i	N	76	36	112	5
	%	68	32		
y	N	23	18	41	2
	%	56	44		
t	N	45	14	59	3
	%	76	24		
e	N	6	33	39	2
	%	15	85		
f	N	32	7	39	2
	%	82	18		
o	N	17	27	44	2
	%	39	61		
w	N	32	47	79	4
	%	41	59		
n	N	43	82	125	6
	%	34	66		
z	N	24	14	38	2
	%	63	37		
q	N	11	12	23	1
	%	48	52		
u	N	47	45	92	4
	%	51	49		
Total	N	1587	560	2147	
	%	74	26		

TOTAL N	1587	560	2147
%	74	26	

Group	Factor	Weight	App/Total	Input&Weight
1:	g	0.048	0.18	0.16
	e	0.658	0.87	0.88
	f	0.168	0.48	0.43
	h	0.412	0.72	0.72
	b	0.022	0.11	0.08
	a	0.303	0.53	0.62
	c	0.032	0.12	0.11
	d	0.279	0.50	0.59
2:	o	0.526	0.73	0.81
	n	0.423	0.76	0.73
3:	j	0.549	0.79	0.82
	p	0.688	0.93	0.89
	d	0.746	0.91	0.92
	k	0.587	0.89	0.84
	s	0.385	0.78	0.70
	l	0.616	0.90	0.86
	g	0.499	0.83	0.79
	v	0.593	0.87	0.85
	h	0.585	0.86	0.84
	a	0.707	0.92	0.90
	b	0.213	0.61	0.50
	c	0.570	0.88	0.83
	m	0.633	0.90	0.87
	r	0.737	0.93	0.91
	i	0.287	0.68	0.60
	y	0.201	0.56	0.49
	t	0.427	0.76	0.74
	e	0.130	0.15	0.36
	f	0.890	0.82	0.97
	o	0.381	0.39	0.70
	w	0.367	0.41	0.68
	n	0.293	0.34	0.61
	z	0.681	0.63	0.89
	q	0.379	0.48	0.70
	u	0.446	0.51	0.75

Cell	Total	App'ns	Expected	Error
hoy	2	0	0.846	1.467
hov	3	3	2.428	0.706
hot	2	2	1.370	0.920
hos	7	6	4.524	1.362
hop	3	2	2.597	1.020
hom	3	3	2.503	0.595
hol	3	3	2.472	0.641
hok	10	4	8.054	10.488

hoj	12	12	9.362	3.381
hoi	10	2	5.404	4.666
hoh	15	15	12.070	3.642
hog	2	1	1.488	0.624
hod	5	5	4.477	0.584
hoc	7	7	5.562	1.810
hob	10	0	4.408	7.884
hoa	3	3	2.628	0.425
hnp	1	1	0.809	0.235
hnm	1	1	0.769	0.301
hnj	5	5	3.504	2.134
hnh	1	1	0.731	0.368
goy	2	0	0.100	0.106
gov	3	0	0.703	0.917
got	5	0	0.676	0.782
gos	2	1	0.232	2.868
gor	1	1	0.371	1.696
gom	1	0	0.266	0.363
gok	3	1	0.689	0.183
goj	4	1	0.814	0.053
goi	4	0	0.312	0.339
goh	4	0	0.915	1.186
gog	7	2	1.210	0.623
god	2	2	0.763	3.244
goc	2	0	0.435	0.557
goa	1	0	0.337	0.508
gnv	1	0	0.168	0.202
gnj	2	0	0.289	0.337
gni	2	1	0.106	7.977
gnh	2	0	0.327	0.391
gnd	1	0	0.289	0.407
foy	3	0	0.524	0.634
fov	1	0	0.550	1.225
fos	6	6	2.070	11.394
fop	2	0	1.300	3.711
fol	4	2	2.297	0.090
fok	2	2	1.088	1.676
foj	5	4	2.529	1.732
foi	1	1	0.253	2.950
foh	3	2	1.628	0.185
fog	2	0	0.911	1.674
fod	9	3	6.406	6.281
fnt	1	1	0.293	2.419
fns	1	0	0.258	0.348
fnm	2	2	0.979	2.086
fnk	2	0	0.881	1.575
fnj	10	2	4.031	1.715
fni	1	1	0.183	4.469
fnh	2	0	0.879	1.567
fnd	2	1	1.239	0.122
fnb	1	1	0.130	6.666
fna	1	1	0.573	0.745
eoy	22	15	14.689	0.020
eov	55	50	50.649	0.105

eot	34	34	29.110	5.711
eos	105	83	87.514	1.398
eor	20	18	19.147	1.610
eop	44	42	41.639	0.058
eom	48	43	44.758	1.022
eol	52	48	48.238	0.016
eok	135	125	124.060	0.088
eoj	56	50	50.778	0.128
eoi	70	54	53.419	0.027
eoh	77	70	70.732	0.093
eog	90	81	79.952	0.123
eod	71	71	68.098	3.026
eoc	55	50	50.257	0.015
eob	47	36	32.126	1.476
eoa	25	24	23.770	0.045
eny	12	8	6.841	0.457
env	13	13	11.503	1.692
ent	17	8	13.551	11.210
ens	37	27	28.403	0.298
enr	24	23	22.482	0.189
enp	17	17	15.655	1.461
enm	23	21	20.725	0.037
enl	19	17	16.993	0.000
enk	32	31	28.229	2.309
enj	32	25	27.687	1.934
eni	24	17	16.324	0.088
enh	70	61	61.715	0.070
eng	16	13	13.441	0.090
end	21	19	19.726	0.441
enc	34	29	29.746	0.150
enb	38	22	22.335	0.012
ena	18	16	16.692	0.395
doz	3	2	2.322	0.198
dow	3	0	1.446	2.792
dou	4	1	2.255	1.602
doq	5	2	2.477	0.182
doo	7	4	3.480	0.154
don	4	3	1.598	2.048
doe	1	1	0.194	4.154
dnu	2	2	0.921	2.344
dnn	1	0	0.305	0.439
coz	3	0	0.676	0.873
cow	9	1	0.660	0.190
cou	6	0	0.594	0.659
coq	1	0	0.077	0.083
coo	8	0	0.620	0.672
con	19	3	1.017	4.083
cof	6	3	3.145	0.014
coe	5	0	0.100	0.102
cnu	1	0	0.068	0.073
cnn	1	0	0.036	0.037
cne	1	0	0.013	0.014
boz	7	0	1.154	1.382
bow	7	1	0.356	1.224

bou	3	0	0.208	0.224
boq	1	0	0.054	0.057
boo	5	0	0.270	0.285
bon	5	2	0.185	18.525
bof	4	1	1.711	0.516
boe	5	0	0.068	0.069
aoz	25	22	19.844	1.135
aow	49	23	25.048	0.343
aou	56	34	33.167	0.051
aoq	16	9	8.394	0.092
aoo	24	13	12.631	0.023
aon	87	32	37.222	1.281
aof	29	28	27.137	0.427
aoe	25	5	5.322	0.025
anw	11	7	4.492	2.366
anu	20	10	9.790	0.009
ann	8	3	2.644	0.072
ane	2	0	0.303	0.357

Total Chi-square = 204.2249
Chi-square/cell = 1.5241
Log likelihood = -908.379

6. **(ing) production, children only, includes style.** Significant factor groups are grammatical form, style, and individual differences.

Number of cells: 94
Application value(s): 1
Total no. of factors: 23

Factor Group 1: Grammatical Form

Key to Factors (Column 1)

n = nouns
v = verbs and complements
a = verbal adjectives
s = something/nothing

Group		Apps	Non-apps	Total	%
1 (2)					
n	N	9	40	49	3
	%	18	82		
v	N	1261	192	1453	87
	%	87	13		
a	N	29	32	61	4
	%	48	52		
s	N	76	29	105	6
	%	72	28		
Total	N	1375	293	1668	
	%	82	18		

Factor Group 2: Style

Key to Factors (Column 1)

o = conversation
n = activity

Group		Apps	Non-apps	Total	%
2					
o	N	990	192	1182	71
	%	84	16		
n	N	385	101	486	29

	%	79	21		
Total	N	1375	293	1668	
	%	82	18		

Factor Group 3: Individual Differences

Key to Factors (See Run #3, Factor Group 3)

Group		Apps	Non-apps	Total	%
3					
j	N	99	27	126	8
	%	79	21		
p	N	62	5	67	4
	%	93	7		
d	N	101	10	111	7
	%	91	9		
k	N	163	21	184	11
	%	89	11		
s	N	123	35	158	9
	%	78	22		
l	N	70	8	78	5
	%	90	10		
g	N	97	20	117	7
	%	83	17		
v	N	66	10	76	5
	%	87	13		
h	N	149	25	174	10
	%	86	14		
a	N	44	4	48	3
	%	92	8		
b	N	59	37	96	6
	%	61	39		
c	N	86	12	98	6
	%	88	12		
m	N	70	8	78	5
	%	90	10		

r	N	42	3	45	3
	%	93	7		
i	N	76	36	112	7
	%	68	32		
y	N	23	18	41	2
	%	56	44		
t	N	45	14	59	4
	%	76	24		
Total	N	1375	293	1668	
	%	82	18		
TOTAL	N	1375	293	1668	
	%	82	18		

Group	Factor	Weight	App/Total	Input&Weight
1:	n	0.032	0.18	0.17
	v	0.563	0.87	0.89
	a	0.120	0.48	0.45
	s	0.314	0.72	0.74
2:	o	0.536	0.84	0.88
	n	0.412	0.79	0.81
3:	j	0.527	0.79	0.87
	p	0.665	0.93	0.92
	d	0.724	0.91	0.94
	k	0.559	0.89	0.89
	s	0.360	0.78	0.77
	l	0.590	0.90	0.90
	g	0.470	0.83	0.84
	v	0.566	0.87	0.89
	h	0.563	0.86	0.89
	a	0.688	0.92	0.93
	b	0.197	0.61	0.60
	c	0.546	0.88	0.88
	m	0.611	0.90	0.91
	r	0.721	0.93	0.94
	i	0.266	0.68	0.69
	y	0.185	0.56	0.58
	t	0.403	0.76	0.80

Cell	Total	App'ns	Expected	Error
voy	22	15	14.839	0.005
vov	55	50	50.739	0.139
vot	34	34	29.252	5.518
vos	105	83	87.895	1.674
vor	20	18	19.188	1.811
vop	44	42	41.707	0.040

vom	48	43	44.877	1.207
vol	52	48	48.325	0.031
vok	135	125	124.295	0.050
voj	56	50	50.987	0.213
voi	70	54	53.752	0.005
voh	77	70	70.986	0.175
vog	90	81	80.103	0.091
vod	71	71	68.163	2.955
voc	55	50	50.419	0.042
vob	47	36	32.531	1.201
voa	25	24	23.820	0.029
vny	12	8	6.679	0.589
vnv	13	13	11.417	1.802
vnt	17	8	13.408	10.324
vns	37	27	28.005	0.148
vnr	24	23	22.433	0.219
vnp	17	17	15.585	1.543
vnm	23	21	20.630	0.064
vnl	19	17	16.881	0.008
vnk	32	31	28.017	2.552
vnj	32	25	27.531	1.667
vni	24	17	16.011	0.183
vnh	70	61	61.410	0.022
vng	16	13	13.290	0.037
vnd	21	19	19.650	0.334
vnc	34	29	29.566	0.083
vnb	38	22	21.913	0.001
vna	18	16	16.640	0.325
soy	2	0	0.848	1.472
sov	3	3	2.426	0.709
sot	2	2	1.373	0.913
sos	7	6	4.523	1.363
sop	3	2	2.598	1.027
som	3	3	2.509	0.588
sol	3	3	2.471	0.642
sok	10	4	8.049	10.440
soj	12	12	9.399	3.321
soi	10	2	5.403	4.663
soh	15	15	12.112	3.577
sog	2	1	1.484	0.612
sod	5	5	4.476	0.586
soc	7	7	5.575	1.790
sob	10	0	4.441	7.989
soa	3	3	2.633	0.418
snp	1	1	0.796	0.255
snm	1	1	0.756	0.323
snj	5	5	3.432	2.284
snh	1	1	0.718	0.394
noy	2	0	0.101	0.107
nov	3	0	0.704	0.921
not	5	0	0.685	0.794
nos	2	1	0.234	2.841
nor	1	1	0.378	1.642
nom	1	0	0.270	0.370

nok	3	1	0.691	0.180
noj	4	1	0.831	0.044
noi	4	0	0.314	0.341
noh	4	0	0.933	1.216
nog	7	2	1.208	0.627
nod	2	2	0.765	3.230
noc	2	0	0.442	0.567
noa	1	0	0.342	0.520
nnv	1	0	0.157	0.186
nnj	2	0	0.274	0.318
nni	2	1	0.098	8.705
nnh	2	0	0.311	0.368
nnd	1	0	0.273	0.375
aoy	3	0	0.537	0.655
aov	1	0	0.556	1.254
aos	6	6	2.107	11.084
aop	2	0	1.314	3.831
aol	4	2	2.323	0.107
aok	2	2	1.100	1.635
aoj	5	4	2.586	1.601
aoi	1	1	0.258	2.869
aoh	3	2	1.663	0.154
aog	2	0	0.920	1.705
aod	9	3	6.451	6.519
ant	1	1	0.282	2.543
ans	1	0	0.247	0.328
anm	2	2	0.957	2.181
ank	2	0	0.851	1.482
anj	10	2	3.936	1.570
ani	1	1	0.174	4.737
anh	2	0	0.859	1.506
and	2	1	1.211	0.093
anb	1	1	0.125	6.971
ana	1	1	0.563	0.776

Total Chi-square = 155.4114
Chi-square/cell = 1.6533
Log likelihood = -640.391

REFERENCES

Andersen, Elaine. 1986. The acquisition of register variation by Anglo-American children. In Bambi Schieffelin and Elinor Ochs (eds.), *Language socialization across cultures.* New York: Cambridge University Press.

Andersen, Elaine. 1990. *Speaking with style: The sociolinguistic skills of children.* New York: Routledge.

Bayley, Robert 1993. Variation in Tejano English: Evidence for variable lexical phonology. Paper presented at the Conference on Language Variety in the South, Auburn University.

Bell, Allan. 1977. The language of radio news in Auckland. Unpublished doctoral dissertation, University of Auckland.

Berko, Jean. 1958. The child's learning of English morphology. *Word*. 14: 150-177.

Bloom, Lois and Margaret Lahey. 1978. *Language development and language disorders*. New York: John Wiley and Sons.

Bowerman, Melissa. 1982. Starting to talk worse: Clues to language acquisition from children's late errors. In S. Strauss (ed.), *U-shaped behavioral growth.* New York: Academic Press.

Bradley, David and Maya Bradley. 1979. Melbourne vowels. University of Melbourne Working Papers in Linguistics, No. 5.

Brown, Roger. 1973. *A first language: The early Stages*. Cambridge, MA: Harvard University Press.

Bybee, Joan and D. I. Slobin. 1982a. Rules and schemes in the development and use of the English past tense. *Language*. 58: 265-289.

Bybee, Joan and D. I. Slobin. 1982b. Why small children cannot change language on their own. In A. Ahlqvist (ed.), *Papers from the 5th International Conference on Historical Linguistics. (Current Issues in Linguistic Theory, Vol. 21, Amsterdam Studies in the theory and history of linguistic science IV.)* Philadelphia/Amsterdam: John Benjamins.

Cazden, Courtney. 1968. The acquisition of noun and verb inflections. *Child Development*. 39: 433-448.

Cedergren, Henrietta and David Sankoff. 1974. Variable rules: Performance as a statistical reflection of competence. *Language*. 50: 333-355.

Crystal, D., Fletcher, P. and Garman, M. 1976. *The grammatical analysis of language disability: A procedure for assessment and remediation.* New York: Elsevier.

Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Cofer, Thomas. 1972. Linguistic variability in a Philadelphia speech community. Unpublished doctoral dissertation, University of Pennsylvania.

Dearborn, Benjamin. 1795. *Columbian grammar*. Boston. As cited in George P. Krapp, *The English language in America*. New York: The Century Co., 1925, Frederick Ungar, 1952.

deVilliers, J. and P. deVilliers. 1985. A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research*. 2(3): 267-278.

Ervin-Tripp, Susan. 1973. Children's sociolinguistic competence and dialect diversity. In A. Dil (ed.), *Language acquisition and communicative choice: Essays by Susan M. Ervin-Tripp.* Stanford: Stanford University Press.

Ervin, Susan and W. R. Miller. 1963. Language development. In H. W. Stevenson (ed.), *Child psychology: The sixty second yearbook of the National Society of the Study of Education, Part I*. Chicago: University of Chicago Press.

Fasold, Ralph. 1972. *Tense marking in Black English*. Arlington, VA: Center for Applied Linguistics.

Ferguson, Charles A. 1977. Babytalk as a simplified register. In Catharine E. Snow and Charles A. Ferguson (eds.), *Talking to children.* Cambridge: Cambridge University Press.

Fischer, John. 1958. Social influence of a linguistic variant. *Word*. 14: 47-56.

Guy, Gregory. 1980. Variation in the group and the individual: The case of final stop deletion. In Labov, William (ed.) *Locating language in time and space*. New York: Academic Press.

Guy, Gregory. 1991. Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change*. 3(1): 1-22.

Guy, Gregory and Sally Boyd. 1990. The development of a morphological class. *Language Variation and Change*. 2(1): 1-18.

Houston, Ann. 1985. Continuity and change in English morphology: The variable (ING). Unpublished Ph.D. dissertation, University of Pennsylvania, Philadelphia.

Kim, John J., Steven Pinker, Alan Prince, and Sandeep Prasada. 1991. Why no mere mortal has ever flown out to center field. *Cognitive Science*. 15: 173-218.

Kiparsky, Paul. 1982. Lexical phonology and morphology. In I. S. Yang (ed.), *Linguistics in the morning calm*. Seoul: Hansen.

Kiparsky, Paul. 1985. Some consequences of lexical phonology. In C. J. Ewens and J. M. Anderson (eds.), *Phonology yearbook 2*. Cambridge: Cambridge University Press.

Klee, Thomas and Martha Deitz Fitzgerald. 1985. The relation between grammatical development and mean length of utterance in morphemes. *Journal of Child Language*. 12: 251-265.

Kovac, Cecil and H. D. Adamson. 1981. Variation theory and first language acquisition. In Sankoff, David and Henrietta Cedergren (eds.) *Variation omnibus*. Edmunton, Alberta: Linguistic Research, Inc.

Krapp, George P. 1925. *The English language in America*. 2 vols. New York: The Century Co., Frederick Ungar, 1952.

Kroch, Anthony. 1978. Toward a theory of social dialect variation. *Language in Society*. 7: 17-36.

Kuczaj, S. A. 1976. -ing, -s, and -ed: A study of the acquisition of certain verb inflections. Unpublished doctoral dissertation. University of Minnesota.

Kuczaj, S. A. 1977. The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior*. 15: 589-600.

Labov, William. 1966. *The social stratification of English in New York City*. Washington, D. C.: Center for Applied Linguistics.

Labov, William. 1967. Some sources of reading problems for Negro speakers of non-standard English. In A. Frazier (ed.), *New directions in elementary English*. Champaign, IL: National Council of Teachers of English.

Labov, William. 1969. Contraction, deletion and inherent variability of the English copula. *Language*. 45: 715-762.

Labov, William. 1970. *The study of non-standard English*. Urbana, IL: National Council of Teachers of English.

Labov, William. 1975. The quantitative study of linguistics structure. In Dahlstedt, Karl-Hampus (ed.), *The Nordic languages and modern linguistics*. Stockholm.

Labov, William. 1980. The social origins of sound change. In Labov, William (ed.) *Locating language in time and space*. New York: Academic Press.

Labov, William. 1984. Field methods on the Project on Linguistic Change and Variation. In Baugh, John and Joel Sherzer (eds.), *Language in use: Readings in sociolinguistics*. Englewood Cliffs, NJ: Prentice-Hall.

Labov, William. 1989a. The child as linguistic historian. *Language Variation and Change*. 1(1): 85-98.

Labov, William. 1989b. Exact description of the speech community: Short A in Philadelphia. In Fasold, Ralph and Deborah Schiffrin (eds.) *Language change and variation.* Amsterdam: John Benjamins.

Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change.* 2(2): 205-254.

Labov, William, P. Cohen, C. Robins, and J. Lewis. 1968. *A study of the non-standard English of Negro and Puerto Rican speakers in New York City.* Cooperative Research Report 3488. Vols. 1 and 2. Philadelphia: U. S. Regional Survey (Linguistics Laboratory, University of Pennsylvania.)

Lakoff, Robin. 1973. Language and women's place. *Language in Society*. 2(1): 45-79.

Maccoby, E. and C. Jacklin. 1974. *The psychology of sex differences.* Stanford: Stanford University Press.

Marchman, Virginia. 1988. Rules and regularities in the acquisition of the English past tense. *Center for research in language newsletter, 2.* April.

Marcus, Gary, Michael Ullman, Steven Pinker, Michelle Hollander, T. John Rosen, and Fei Xu. 1990. *Overregularization*. Occasional paper #41. Cambridge, MA: MIT, The Center for Cognitive Science.

Menyuk, Paula. 1963. A preliminary evaluation of grammatical capacity in children. *Journal of Verbal Learning and Verbal Behavior*. 2: 429-439.

Menyuk, Paula. 1964. Alternation of rules in children's grammar. *Journal of Verbal Learning and Verbal Behavior*. 3: 480-488.

Menyuk, Paula. 1969. *Sentences children use*. Cambridge, MA: MIT Press.

Menyuk, Paula. 1977. *Language and maturation*. Cambridge, MA: MIT Press.

Miller, Jon F. and Robin S. Chapman. 1981. The relation between age and mean length of utterance in morphemes. *Journal of Speech and Hearing Research.* 24: 154-161.

Miller, Peggy and Catharine Garvey. 1984. Mother-baby role play: Its origins in social support. In I. Bretherton (ed.) *Symbolic play: The representation of social understanding.* New York: Academic Press.

Mohanan, K. P. 1986. *The theory of lexical phonology.* Dordrecht: Reidel.

Moore, S., S. Meech, and H. Whitehall. 1935. *Middle English dialect characteristics and dialect boundaries.* Ann Arbor: University of Michigan Language and Literature Series.

Nelson, Katharine. 1973. Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development,* 38 (Serial No. 149).

Nelson, Katharine. 1975. The nominal shift in semantic-syntactic development. Cognitive Psychology, 7: 461-479.

Neu, Helene. 1980. Ranking of constraints on /t,d/ deletion in American English: A statistical analysis. In Labov, William (ed.) *Locating language in time and space.* New York: Academic Press.

Pinker, Steven and Alan Prince. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition.* 73-193.

Plunkett, Kim and Virginia Marchman. 1991. U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition.* 38: 43-102.

Purcell, April. 1984. Code shifting Hawaiian style: Children's accommodation as a decreolizing continuum. *International Journal of the Sociology of Language.* 46: 71-86.

Rand, David and David Sankoff. 1990. *Goldvarb, Version 2: A variable rule application for the Macintosh.* University of Montreal.

Reid, E. 1978. Social and stylistic variation in the speech of children. In Trudgill, Peter (ed.) *Sociolinguistic patterns in British English*. Baltimore: University Park Press.

Roberts, Julie and William Labov. 1993. Acquisition of a dialect. Paper presented at the New Ways of Analyzing Variation (NWAV) Conference, October, Ann Arbor, MI.

Romaine, Suzanne. 1978. Postvocalic /r/ in Scottish English: Sound change in progress? In Trudgill, Peter (ed.) *Sociolinguistic patterns in British English*. Baltimore: University Park Press.

Rumelhart, D. E. and J. L. McClelland. 1986. On learning the past tenses of English verbs. In J. L. McClelland, D. E. Rumelhart, and the PDP Research Group (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition: Vol.2. Psychological and biological models.* Cambridge, MA: Bradford Books/MIT Press.

Sankoff, Gillian and Suzanne Laberge. 1973 [1980]. On the acquisition of native speakers by a language. *Kivung*, 6, 32-47. Also in Sankoff, Gillian, *The social life of language*. Philadelphia, PA: University of Pennsylvania Press.

Santa Ana, Otto. 1991. Chicano English evidence for the exponential hypothesis. Paper presented at NWAV conference, 1991, Georgetown University.

Scollon, Ron and Suzanne B. K. Scollon. 1981. *Narrative, literacy and face in interethnic communication.* Norwood, NJ: Ablex.

Shatz, M. and R. Gelman. 1973. The development of communication skills: Modifications in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development, no. 38 (Serial No. 152).*

Shuy, Roger, Walt Wolfram, and William K. Riley. 1968. *A study of social dialects in Detroit. (Final Report, Project 6-1347).* Washington, D. C.: Office of Education.

Slobin, D. 1971. On the learning of morphological rules: A reply to Palermo and Eberhart. In D. I. Slobin (ed.), *The ontogenesis of grammar: A theoretical symposium*. New York: Academic Press.

Stemberger, J. P. 1989. The acquisition of morphology: Analysis of a symbolic model of language acquisition. Unpublished manuscript, Department of Linguistics, University of Minnesota.

Strauss, S. 1982. *U-shaped behavioral growth.* New York: Academic Press.

Summerlin, Nan Jo Corbitt. 1972. A dialect study: Affective parameters in the deletion and substitution of consonants in the deep South. Unpublished doctoral dissertation. Florida State University.

Sutton-Smith, Brian. 1971. A syntax for play and games. In R. E. Herron and Brian Sutton-Smith (eds.) *Child's play*. New York: John Wiley and Sons.

Trudgill, Peter. 1974. Sex, covert prestige, and linguistic change in an urban British English in Norwich. *Language in Society*.

van Hout, Roeland. 1989. Analyzing language variation by multiple regression: The case of (t) deletion in a Dutch dialect. In M.E.H. Schouten and P.Th. van Reenen, (eds.) *New methods in dialectology*. Dordrect, Holland: Foris.

Wald, Benji and Timothy Shopen. 1985. A researcher's guide to the sociolinguistic variable (ING). In Virginia Clark, Paul Eschholz, and Alfred Rosa (eds.), *Language: Introductory readings.* New York: St. Martin's Press.

Watson-Gegeo, Karen and David Gegeo. 1986. The social world of the Kwaraa 'ae children: Acquiring language and values. In J. Cook-Gumperz, W. Corsaro, and J. Streek (eds.) *Children's worlds, children's language*. The Hague: Mouton.

Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. In W. Lehmann and J. Malkiel (eds.), *Directions for historical linguistics*. Austin: University of Texas Press.

Wolf, M. and D. Gow. 1986. A longitudinal investigation of gender differences in language and reading development. *First Language*, 6, 81-110.

Wolfram, Walt. 1969. *A sociolinguistic description of Detroit Negro speech*. Washington, D. C.: Center for Applied Linguistics.

Wolfram, Walt. 1972. Overlapping influence and linguistic assimilation in second generation Puerto Rican English. In David Smith and Roger Shuy (eds.), *Sociolinguistic cross-cultural analysis.* Washington, D. C.: Georgetown University Press.

Wolfram, Walt and Ralph Fasold. (1974) *The study of social dialects in American English*. Englewood Cliffs, NJ: Prentice-Hall.

Wolfram, Walt and Donna Christian. 1976. *Appalachian speech*. Washington, D. C.: Center for Applied Linguistics.

Woods, Howard. 1979. A socio-dialectal survey of the English spoken in Ottawa: A study of sociolinguistic and stylistic variation. Unpublished doctoral dissertation, University of Ottawa.

Wyld, Henry. 1936. *A history of modern colloquial English.* Oxford: Basil Blackwell.