

Acquisition of sociolinguistic variables in Spanish: Do children acquire individual lexical forms or variable rules?

Manuel Díaz-Campos

1. Introduction

Previous research in the area of phonological variation has focused on describing internal and external constraints in the speech of adult speakers. These previous investigations have contributed to our understanding of the role played by different groups within the speech community in the process of language change. The study of variation in child language was not taken into consideration for a long period of time in sociolinguistic studies. Pioneer work on the acquisition of variation in child phonology (Labov 1964) proposes that development of stylistic variation probably starts when individuals are 14 years old under the influence of wider contacts with peers beyond the neighborhood or high school. More recently this idea has been challenged by some scholars who have conducted research on the acquisition of variable phonology in English, French, and Spanish (Roberts and Labov 1995; Roberts 1994, 1997a, 1997b; Chevrot, Beaud, and Varga 2000; Díaz-Campos 2001). The assumption in the work of Roberts and Labov 1995; Roberts 1994, 1997a, 1997b, and Díaz-Campos 2001 is that the acquisition of variable phonology entails the encoding of a variable rule. According to Labov (1972), variable rules are based on generative phonological rules with the ingredient of incorporating the probability of application of them when linguistic and social constraints are satisfied. Nonetheless, Chevrot, Beaud, and Varga (2000:295) suggest that children tend to copy adult surface forms instead of acquiring a variable rule. This suggestion that children copy lexical forms is consistent with Bybee's (2001) usage-based model of phonology in which linguistic regularities are not expressed as rules, but rather as schemas. This means that speakers discover generalizations about linguistic units and create a series of connections based on similarities among them.

The purpose of this investigation is to examine whether children acquire sociolinguistically variable phonology as a rule or as a case-by-case copy-

ing of adults' surface forms. The analysis is based on previous work studying the acquisition of intervocalic /d/ in Spanish-speaking children from Venezuela (Díaz-Campos 2001).

The sections included in this paper will be organized as follows. In the next section, a discussion of the previous work is presented focusing on the issues of whether children acquire individual lexical forms or variable rules. Section 3 describes the corpus and methodology used in the present investigation. Section 4 presents the results. Finally, the summary and conclusions are presented in Section 5.

2. Previous research

Investigations examining acquisition of variable phonology are not very common in sociolinguistic literature. Nonetheless, one pioneer study (Labov 1964) and recent studies (Roberts 1994; Roberts and Labov 1995; Roberts 1997a and 1997b; Chevrot, Beaud, and Varga 2000; Díaz-Campos 2001) have explored the subject in order to describe the role of children within the speech community.

One of the first investigations examining acquisition of sociolinguistic variables is that of Labov (1964: 92). This scholar is concerned with the acquisition of "the full range" of spoken English, including standard varieties as well as regional or vernacular varieties of English. Specifically, Labov studies the fortition of /θ/ in words such as *thing*, *think*, *three*, etc. and r-deletion word-finally and in preconsonantal position (e.g. *car*, *board*). Labov's investigation proposes that children from about 5 to 12 years old acquire the vernacular variety under the influence of peers in school and in the neighborhood. Stylistic variation, according to Labov's proposal, would be acquired around age 14 under the influence of wider contact outside the immediate community.

This perspective according to which stylistic variation is acquired during teen years has been changed by current research in the field. Sociolinguistic variables appear in the speech of children as young as 3;2 (see Roberts and Labov 1995, Roberts 1994, 1997a, 1997b). Roberts and Labov (1995) study the acquisition of English short-a, focusing on speech samples from children born in Philadelphia. The results show that 3- and 4-year-old children were acquiring the norms of the speech community in regard to the short-a pattern. Children were acquiring this sociolinguistic variable even in environments of lexical change in progress (short-a before /l/ and before inter-vocalic /n/). According to Roberts and Labov, these findings

show that during the preschool period children are learning local grammatical rules and variable rules.

Roberts (1997a) also examines the acquisition of short-a in Philadelphia children from Philadelphia. She compares the results reported by Roberts and Labov (1995) with her own. The variables are: 1) The fronting and raising of /a/ in *crow*, *south*; 2) the raising of the nucleus in *rate*; 3) the backing of long /ay/ before /t/ in *fight*, *right*, *mice*. Roberts' results show that children in Philadelphia progress in learning Philadelphia's vowel system between 3 years 2 months old to 4 years 11 months old. Furthermore, even though the conditioning of /ey/ is more complex, all children acquire it by age 4. Based on the findings described above, Roberts concludes that preschool years are learning their local dialect. This change in progress seems to be influenced by the speech of their parents; and 3) it is the female-dominated variety that advanced in early language acquisition.

Roberts (1994 and 1997b) examines the acquisition of word-final consonant clusters in children from Philadelphia. /d/ deletion is a well documented English feature. Roberts finds it an interesting phenomenon to study in Philadelphia. Roberts' results reveal that children acquire patterns of /t/ and /d/ deletion in the Philadelphia dialect. Roberts' findings come as an indication that children simultaneously acquire simple and complex grammatical forms.

In the case of Spanish, Díaz-Campos (2001) examines the acquisition of sociolinguistic variables in speech samples from Philadelphia children. Two linguistic variables are examined: 1) /d/ deletion, and 2) syllable-final /r/. Díaz-Campos (2001) finds that children at an early age (i.e. 4 years 5 months old), children in Philadelphia acquire a sociolinguistic value that is similar to that of Philadelphia. Children acquire first the sociolect of their parents. Through regular exposure to other systems, they build up their repertoires in their speech, but also assign values to them.

There are many points in common in the findings of Roberts and Labov (1995); Roberts (1997a), (1997b), and Díaz-Campos (2001), but the one we need to point out for the present

show that during the preschool period children are able to acquire categorical grammatical rules and variable rules.

Roberts (1997a) also examines the acquisition of variable phonology in children from Philadelphia. She compares her findings with previous results reported by Roberts and Labov (1995). Roberts studies three linguistic variables: 1) The fronting and raising of the nucleus of /aw/, as in *cow*, *crown*, *south*; 2) the raising of the nucleus of checked long /ey/, as in *cake* and *rate*; 3) the backing of long /ay/ before voiceless final obstruents, as in *fight*, *right*, *mice*. Roberts' results show that all children were making progress in learning Philadelphia's vowel system. Children ranged in age from 3 years 2 months old to 4 years 11 months had acquired the fronting of /aw/. Furthermore, even though the conditioning of the raising of the nucleus of /ey/ is more complex, all children acquired this change in progress as well. Based on the findings described above, Roberts claims that 1) children in preschool years are learning their local dialect; 2) the acquisition of sound change in progress seems to be influenced by the dialect background of their parents; and 3) it is the female-dominated sound changes that are advanced in early language acquisition.

Roberts (1994 and 1997b) examines the deletion of final /t/ and /d/ in word-final consonant clusters in children from Philadelphia. Since /t/ and /d/ deletion is a well documented English variable rule, the author considers it an interesting phenomenon to study in children's speech in order to understand the acquisition of variation. As in Roberts' previous work, the results reveal that children acquire patterns of variation at an early age. The author found that children were learning the internal constraints that govern /t/ and /d/ deletion in the Philadelphia dialect. Roberts interprets this outcome as an indication that children simultaneously learn patterns of variation and complex grammatical forms.

In the case of Spanish, Díaz-Campos 2001 has studied the acquisition of sociolinguistic variables in speech samples from 30 Venezuelan preschool children. Two linguistic variables are examined in his work: 1) intervocalic /d/, and 2) syllable-final /r/. Díaz-Campos' results reveal that, from a very early age (i.e. 4 years 5 months old), children begin to use variable phonology with a sociolinguistic value that is similar to that of the adult model. Children acquire first the sociolect of their immediate community, but, with regular exposure to other systems, they begin not only incorporating new repertoires in their speech, but also assigning social value to them.

There are many points in common in the work of Roberts (1994); Roberts and Labov (1995); Roberts (1997a), (1997b), and Díaz-Campos (2001), but the one we need to point out for the present paper is the theoretical as-

sumption that children are acquiring variable rules. Even though current models in generative phonology have abandoned the idea of a rule-based system of phonology, sociolinguistic studies still use this notion for describing social variation. According to Cedergren and Sankoff (1974), variable rules formalize those proposals in sociolinguistic research that examine variability as a core element of linguistic competence. Variable rules incorporate the statistically predicted frequency of its application as an integral part of its structural description. If we take these theoretical considerations into account, we would expect the application of the variable rule any time the structural description proposed is met.

Contrary to the variable rule perspective in the case of acquisition of sociolinguistic variables, Chevrot, Beaud, and Varga (2000) sustain that children copy adult surface lexical forms instead of encoding a variable rule. These scholars study post-consonantal word-final /R/ in French speaking children. Chevrot, Beaud, and Varga (2000) include in their study two age groups children: 6 to 7 year-old and 10 to 12 year-old. According to their findings, they locate variability within the "lexical knowledge internalized for each word" (Chevrot, Beaud, and Varga 2000: 315). The location of variability at the lexical level allows explaining the cognitive nature of speech planning as well as the relationship of variable phonology and social constraints. Chevrot, Beaud, and Varga (2000) point out that cognitive constraints are related to speech production phenomena such as simplification of complex structures. Social factors affecting variability include evaluation of social variants, and the relationship between variants and speech register. In Chevrot, Beaud, and Varga's (2000: 315) words:

This conception allows us to draw together two opposing tendencies that has divided sociolinguistic approaches to variation: an interactional approach, which supposes that speakers are free to manipulate variants in order to achieve communicative aims, versus a variationist approach, which supposes that speakers' linguistic behavior is constrained by linguistic and sociological factors.

As can be seen in the review of the previous research, the variable rule hypothesis would be triggered any time factors favoring variation increase the probability of application of the rule when the phonetic context is met. If we follow the reasoning behind this proposal, we would expect to find alternation in children's speech in different age groups. In other words, if we think in terms of variable rules, we would expect that children regardless of age will use such rules in the corresponding phonetic context any time internal and external constraints favoring its application are met. In

the specific case analyzed here, we would expect to find alternation of words such as [kantáðo], [kantáðo] *lado*, 'side'; etc.

Alternatively, if children acquire variable rules, we would not find alternation at the beginning of the acquisition of individual lexical forms and the incorporation of new forms where stylistic variation begins to emerge. If we find alternation, we would need to examine frequency of use (Bybee (2001)). Following Bybee's proposal, we would expect frequent tokens as opposed to infrequent ones.

3. Methodology

3.1. Corpus

The corpus of the present investigation consists of the monolingual Spanish-speaking children's speech recordings are part of a larger corpus, *Competencia narrativa escolar*¹ 'Narrative competence of school children' by Dr. Martha Shiro, professor of the Master's Program at Universidad Central de Venezuela.

Each interview lasts approximately 45 minutes. The style of the conversation can be described as semi-structured. The first session included four different parts. The first part was with the children including questions such as (1) How did you celebrate your birthday? (2) How many brothers? (3) How many sisters? (4) Do you play with your brothers and sisters with them? (5) What does your father do for a living? (6) What does your mother do for a living? (7) Does your father wake you up at night? (8) Does your father tell you a story? (9) Can you tell that story? (10) What is your favorite TV program about? (11) What is your favorite TV program about? There were several other questions.

The second part of the interview included questions where the interviewer was trying to obtain a story that the child was emotionally involved. Labov (1972) suggests that the speaker's emotional reaction are more likely to be elicited in a classical Labovian interview includes topics that are emotionally charged in order to initiate the emotional involvement.

le rules. Even though current
 oned the idea of a rule-based
 s still use this notion for de-
 gren and Sankoff (1974), vari-
 olinguistic research that exam-
 ic competence. Variable rules
 ncy of its application as an in-
 take these theoretical consid-
 application of the variable rule
 s met.

n the case of acquisition of so-
 Varga (2000) sustain that chil-
 d of encoding a variable rule.
 l-final /R/ in French speaking
 include in their study two age
 2 year-old. According to their
 lexical knowledge internalized
 a 2000: 315). The location of
 ining the cognitive nature of
 of variable phonology and so-
 (2000) point out that cognitive
 phenomena such as simplifica-
 affecting variability include
 ionship between variants and
 a's (2000: 315) words:

two opposing tendencies that
 variation: an interactional ap-
 e to manipulate variants in or-
 a variationist approach, which
 constrained by linguistic and

ous research, the variable rule
 ors favoring variation increase
 en the phonetic context is met.
 osal, we would expect to find
 age groups. In other words, if
 ld expect that children regard-
 ponding phonetic context any
 ing its application are met. In

the specific case analyzed here, we would expect to find alternant produc-
 tion of words such as [kantáðo], [kantáo], *cantado* 'sung'; [laðo], [láo],
lado, 'side'; etc.

Alternatively, if children acquire variation in a case-by-case fashion we
 would not find alternation at the beginning stages, but the production of
 individual lexical forms and the incorporation of new forms at later stages
 where stylistic variation begins to emerge in their speech. Even if we find
 alternation, we would need to examine frequency effects as suggested by
 Bybee (2001). Following Bybee's proposal, we would expect variability in
 frequent tokens as opposed to infrequent ones.

3. Methodology

3.1. Corpus

The corpus of the present investigation consists of recordings in which 30
 monolingual Spanish-speaking children were participants. These 30 re-
 cordings are part of a larger corpus, *Competencia narrativa en niños de
 edad escolar*¹ 'Narrative competence of school-age children' collected by
 Dr. Martha Shiro, professor of the Master's Program in Linguistics at the
 Universidad Central de Venezuela.

Each interview lasts approximately 45 minutes to 1 hour. The speech
 style of the conversation can be described as spontaneous. The recording
 session included four different parts. The first one included a conversation
 with the children including questions such as: (1) how old are you? (2)
 How did you celebrate your birthday? (3) Do you have brothers and sis-
 ters? (4) Do you play with your brothers and sisters? (5) Do you get along
 with them? (6) What does your father do for a living? (7) What does your
 mother do for a living? (8) Does your father or your mother read stories to
 you at night? (9) Can you tell that story? (10) Who do you play with at
 school? (11) What is your favorite TV program? (12) What is that TV pro-
 gram about? There were several other questions with the same characteris-
 tics.

The second part of the interview included a trigger situation in which
 the interviewer was trying to obtain a story where the children were emo-
 tionally involved. Labov (1972) suggests that topics triggering the
 speaker's emotional reaction are more likely to produce casual speech. The
 classical Labovian interview includes topics such as the danger of death in
 order to initiate the emotional involvement of the speaker during the con-

versation. Labov points out that talking about accidents, sickness, and natural disasters may be appropriate to generate the conditions for emotional speech. The topics suggested by Labov (1972) were adapted to the interviews of the children in this research. The interviewer shared a personal experience with the child as follows: "One day I was in the kitchen serving myself a glass of Coke and I did not hold the glass firmly. The glass fell down and broke, hurting my feet. Have you ever had something like that happen to you?" Another trigger situation included the following: "One day I was very sick and had to go to the doctor. Have you ever visited the doctor? What was that experience like?" The trigger situations worked very well for almost all the children. They were able to narrate a situation and to express their emotional reactions.

The third part of the interview was play activity with toys representing the characters from the movie *Aladdin*. The characters from *The Flintstones* were also used during the interview. The children were able to narrate their own version of the movie or to create new situations playing with the characters. Lower- and upper-class speakers were very familiar with the characters from both *Aladdin* and *The Flintstones*.

Finally, the last part of the interview included a storytelling activity using the tale *La gallinita* 'The little hen'. The child had to retell the story with the pictures of the characters and events in the book. All the situations contained during the interview created a positive environment for obtaining a large amount of discourse suitable for performing a quantitative analysis.

3.2. Analyses

In order to test the variable rule and the case-by-case hypotheses, we have examined the number of times a token was produced and whether there was alternation in its production. If we find alternation we would have to observe whether this variability affects all possible targets where the phonetic as well as internal and external constraints are met regardless of age group if we assume an early acquisition of variability. If this were the case, then we would have enough evidence to support the variable rule hypothesis.

Contrastively, the case-by-case hypothesis would be supported if we do not find alternation at the beginning stages. As explained before, even if we find alternation we will have to analyze whether such variability always applies given the same conditions (e.g., similar internal and external factors affecting the linguistic variables). Following the ideas of Bybee's (2001) usage-based model of phonology, it would be possible to think that fre-

quency may play a role in the spread of forms. Frequency effects were measured. The cases analyzed were divided according to frequency. Second, the *Frequency dictionary of Spanish* (Rodríguez 1964) was used to determine the frequency of each token.

GoldVarb 2001, a linguistic software package, was used for the study of sociolinguistic variation. The package was used for the analysis of the data. Five factor groups were defined for examining both the variable rule hypothesis and the case-by-case description of each factor group is presented below.

- 1) *Factor group 1* (Alternation/Retention) This factor group has been taken as the dependent variable to test whether each of the tokens analyzed in the study explained above, one way to test the case-by-case hypotheses is to examine the production of intervocalic /d/. The case-by-case: 1) alternation, and 2) retention.
- 2) *Factor group 2* (Deletion/Retention) This factor group measures whether /d/ is deleted or retained (e.g., /d/ deletion or retention). Tokens with no /d/ were not found. Tokens with no /d/ are: 1) cases of /d/ deletion or retention.
- 3) *Factor group 3* (Dictionary Frequency) This factor group examines frequency effects by consulting the *Frequency dictionary of Spanish* (Rodríguez 1964). It consists of 5,024 words coming from different genres: dramatic literature (plays), fictional literature (novels), sayistic literature (essays, newspaper articles), scientific literature (medicine, engineering), and popular literature (dailies, weeklies, magazines). The factor group taken into account: 1) high frequency (general usage), 2) mid frequency (intermediate usage), and 3) low frequency (specialized usage).
- 4) *Factor group 4* (Corpus Frequency) This factor group considers the frequency of the most frequent words contained in the corpus. The most frequent words containing /d/ and /t/ are having high frequency. The least frequent words containing /d/ and /t/ are having low frequency. Included in this factor group are the most frequent words containing /d/ and /t/.

quency may play a role in the spreading of variation in individual lexical forms. Frequency effects were measured in two different ways. First, the cases analyzed were divided according to their frequency in the corpus. Second, the *Frequency dictionary of Spanish words*² (Jullian and Chang-Rodríguez 1964) was used to determine the coefficient of general usage for each token.

GoldVarb 2001, a linguistic software program specifically designed for the study of sociolinguistic variation, was used for performing a statistical analysis of the data. Five factor groups were included in the analysis for examining both the variable rule and the case-by-case hypotheses. A description of each factor group is presented below:

- 1) *Factor group 1* (Alternation vs. no alternation): This Factor group has been taken as the dependent variable. It takes into consideration whether each of the tokens was produced variably or not. As explained above, one way to examine both the variable rule and the case-by-case hypotheses is considering whether there is variably in the production of intervocalic /d/. Two factors are included in this case: 1) alternation, and 2) no alternation.
- 2) *Factor group 2* (Deletion/retention in non-alternating cases): This Factor group measures what is the predominant production tendency (e.g., /d/ deletion or retention) in the tokens in which alternation was not found. Tokens with no alternation were classified in two categories: 1) cases of /d/ deletion, and 2) cases of /d/ retention.
- 3) *Factor group 3* (Dictionary frequency³): In this case, we examine frequency effects by considering the coefficient of general usage proposed by Jullian and Chang-Rodríguez (1964) in their *Frequency dictionary of Spanish words*. This frequency dictionary consists of 5,024 words coming from 5 different sources including dramatic literature (plays), fictional literature (novels, short stories), essayistic literature (essays, memoirs, correspondence, etc.), technical literature (medicine, engineering, physics, etc.), and journalistic literature (dailies, weeklies, monthlies). Three levels of frequency were taken into account: 1) high frequency (101 or higher coefficient of general usage), 2) mid frequency (50-100 coefficient of general usage), and 3) low frequency (0-49 coefficient of general usage).
- 4) *Factor group 4* (Corpus frequency): This Factor group takes into consideration the frequency of the tokens within the corpus. The 20 most frequent words containing an intervocalic /d/ were classified as having high frequency. The rest of the tokens were considered as having low frequency. Including this measurement will allow com-

paring whether the tendencies found while examining the Julliard and Chang-Rodríguez (1964) frequency dictionary pattern in the same way.

- 5) *Factor group 5 (Age)*: In this case, we distinguish 6 different groups: (1) 42- to 47-month-olds, (2) 48- to 53-month-olds, (3) 54- to 59-month-olds, (4) 60- to 65-month-olds, and (5) 66- to 71-month-olds. The motivation for having 6 groups is to track with more detail phonological acquisition at different stages of development. Previous work related to the acquisition of sociolinguistic variables in Spanish has shown that this age division is important for explaining changes in children's speech (see Díaz-Campos 2001).

4. Results and discussion

4.1. Results

This section is divided in two parts. The first one presents the 3 factor groups selected as significant by GoldVarb 2001. The second one discusses the findings in order to answer whether sociolinguistic variables are acquired in a case-by-case fashion or as variable rules.

GoldVarb 2001 allows for clarification of how a group of internal and external constraints are related to the dependent variable. In the particular case of this investigation, factor group 1 (alternation vs. not alternation in the production of intervocalic /d/) is the dependent variable. The application value is factor *alternation* since examining whether or not there is variability across children of different age groups will allow testing the hypotheses proposed at the end of Section 2.2. GoldVarb reveals the probabilistic weight of each one of the constraints not only in relation to a given variant but also the overall effect of significant factors regarding that variant. A weight greater than .500 favors the application value and a lesser probability disfavors it.

The factors found to be statistically significant are presented in Table 1⁴ in order of selection. This order of selection is crucially important because it indicates the degree of impact of the factor group selected on the dependent variable from most important to least important. According to this explanation, factor group 3, which measures frequency effects considering the coefficient of general usage proposed by Julliard and Chang-Rodríguez (1964), has the greatest impact on alternant tokens (a discussion of this result is given below). Factor group 4 (frequency within the corpus) is the

next in the hierarchy. The next constraint is (alternation vs. not alternation) (sustant). The original analysis also included the degree of /d/ deletion or retention in non-dependent constraint contains *knockout factor*. It is not possible to be able to perform a probabilistic analysis. If there are empty cells and this condition does not allow for a probabilistic analysis.

Table 1: Factor groups selected in the acquisition of sociolinguistic variables analysis (input probability = 0.041 (6))

Factor group	Factors	No of cases
Dictionary frequency	High	27/58
	Mid	3/15
	Low	22/174
Corpus frequency	High	44/80
	Low	19/245
Age	42-47 months	19/85
	48-53 months	7/41
	54-59 months	13/58
	60-65 months	13/88
	66-71 months	11/72

The factor group selected as the most important is frequency. The results reveal that high frequency words show a borderline tendency according to the production of intervocalic /d/ with a frequency of lower frequency words, the findings indicate that the production of intervocalic /d/ is neither favored nor disfavored. Dictionary frequency is the one having the greatest impact. The production of sociolinguistic variables is more likely to occur with high frequency tokens at the beginning stages. Lower frequency words in the production of intervocalic /d/, so we would expect that the production would spread to this class of tokens at the beginning stages. In the following section, we discuss in more detail the findings and whether they support the case-by-case or the variable rule hypothesis.

The second factor group selected is corpus frequency. This factor is very important because it is consistent with the findings of high frequency tokens.

while examining the Jullian and Chang-Rodriguez dictionary pattern in the

distinguish 6 different groups: (1) 53-month-olds, (2) 54- to 59-month-olds, (3) 54- to 59-month-olds, (4) 60- to 65-month-olds, and (5) 66- to 71-month-olds. This study tracks with more detail the phonological stages of development. Previous studies of sociolinguistic variables in Spanish have been important for explaining changes in the acquisition of /d/ (Jullian and Chang-Rodriguez 2001).

The first one presents the 3 factor groups selected in the acquisition of sociolinguistic variables are age, dictionary frequency and corpus frequency.

The second one presents the 3 factor groups selected in the acquisition of sociolinguistic variables are age, dictionary frequency and corpus frequency. The third one presents the 3 factor groups selected in the acquisition of sociolinguistic variables are age, dictionary frequency and corpus frequency. The fourth one presents the 3 factor groups selected in the acquisition of sociolinguistic variables are age, dictionary frequency and corpus frequency. The fifth one presents the 3 factor groups selected in the acquisition of sociolinguistic variables are age, dictionary frequency and corpus frequency.

The factor groups selected in the acquisition of sociolinguistic variables are age, dictionary frequency and corpus frequency. The factor groups selected in the acquisition of sociolinguistic variables are age, dictionary frequency and corpus frequency. The factor groups selected in the acquisition of sociolinguistic variables are age, dictionary frequency and corpus frequency. The factor groups selected in the acquisition of sociolinguistic variables are age, dictionary frequency and corpus frequency.

next in the hierarchy. The next constraint is factor group 5 (age of the consultant). The original analysis also included factor group 2, which considers degree of /d/ deletion or retention in non-alternating tokens. Since this independent constraint contains *knockout factors*, it was discarded in order to be able to perform a probabilistic analysis. A *knockout factor* occurs when there are empty cells and this condition does not allow GoldVarb to run the probabilistic analysis.

Table 1: Factor groups selected in the acquisition of sociolinguistic variable analysis (input probability = 0.041 (63/344))

Factor group	Factors	No of cases	%	Weight
Dictionary frequency	High	27/58	46	.775
	Mid	3/15	20	.515
	Low	22/174	12	.397
Corpus frequency	High	44/80	55	.991
	Low	19/245	7	.196
Age	42-47 months	19/85	22	.972
	48-53 months	7/41	17	.093
	54-59 months	13/58	22	.288
	60-65 months	13/88	14	.312
	66-71 months	11/72	15	.235

The factor group selected as the most significant is dictionary frequency. The results reveal that high frequency tokens trigger alternation in the production of intervocalic /d/ with a weight of .775. Mid frequency words show a borderline tendency according to which alternation in the production of intervocalic /d/ is neither favored nor disfavored. In the case of lower frequency words, the findings indicate that they disfavor variability in the production of intervocalic /d/. The fact that the factor group dictionary frequency is the one having the greatest impact on the variable pronunciation of intervocalic /d/ is a piece of evidence revealing that acquisition of sociolinguistic variables is more likely to happen in high frequency tokens at the beginning stages. Lower frequency tokens disfavor alternation in the production of intervocalic /d/, so we would expect that stylistic variation would spread to this class of tokens at a more advanced stage. In the following section, we discuss in more detail these findings and explain whether they support the case-by-case or the variable rule hypotheses.

The second factor group selected is corpus frequency. This result is also very important because it is consistent with the factor group dictionary frequency. The weight for high frequency tokens within the corpus is .991,

which reveals that this factor highly favors variable production of intervocalic /d/. In the case of lower frequency tokens, alternation in the production of intervocalic /d/ is disfavored. Among the most frequent tokens within the corpus (these tokens also have a high frequency according to Jullian and Chang-Rodríguez 1964) we find *de* 'of/from', *lado* 'side', *donde* 'where', *todavía* 'yet', etc. In all of these tokens, alternation in the production of intervocalic /d/ is the predominant tendency.

Age is the next most important factor group according to the statistical analysis. The factor with the highest weight is 42- to 47- month old children, at .972. This result indicates that the youngest children favor variable pronunciation. In the rest of the age groups, the findings show that alternation in the pronunciation of intervocalic /d/ is disfavored. Is variability in the production of intervocalic /d/ a phenomenon relating to the development of the phonology of the youngest group? How can we explain these tendencies? What is the favored pronunciation in non-alternating cases? In a previous study examining deletion of intervocalic /d/ in this same group of children (see Díaz-Campos 2001), it was determined that the variable pronunciation of intervocalic /d/ is not developmental, but sociolinguistic in nature. Children show in their speech from a very early age the patterns of variation of their immediate community. The fact that older children do not favor variability is reflecting the effect of the variety of language spoken at school where variability is less likely to happen and retention of intervocalic /d/ is the predominant tendency. Retention is favored in the speech of upper class children. Older lower class children begin to retain more intervocalic /d/ in their pronunciation once they are exposed to the school variety of language. This explanation is corroborated by the results of the present analysis according to which in non-alternating cases the predominant pronunciation favors retention of intervocalic /d/ (236 tokens out of 281). In any case, it is important to point out that even though variability is not favored by older children, these age groups did show alternation in their pronunciation of intervocalic /d/ in highly frequent tokens as discussed above.

4.2. Discussion

Since the main concern of this investigation is to determine whether children acquire sociolinguistic variation in a case-by-case fashion or acquire variable rules, we turn our attention to a discussion of the hypotheses presented in Section 2. The statistical results indicate that the two factor

groups, dictionary frequency as well as role for predicting variable pronunciation evidence will be an important element in that we will develop in the next section.

According to the hypothesis presented to answer whether we find variability in different age groups. Alternating pronunciation could be an indication of acquisition of variable rules. We will explore further pieces of evidence to support this hypothesis. Figure 1 shows the patterns of alternation in different age groups.

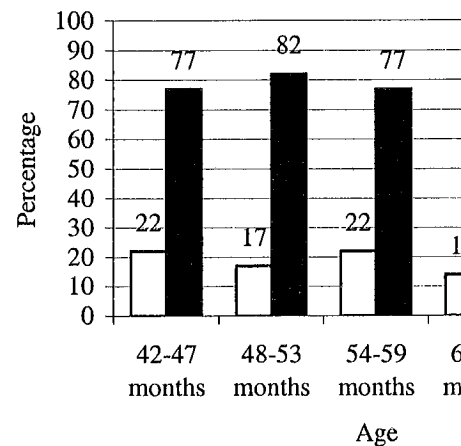


Figure 1. Pattern of alternation in the production of intervocalic /d/ across age groups

Because of the probabilistic analysis that 42- to 47- month old children are more likely to alternate the production of intervocalic /d/. However, figure 1 shows that the 48-53 month old children show the highest percentage of alternation in frequent tokens regarding alternating pronunciation. This fact is reflected in Figure 1. This result supports the hypothesis that the variable rule analysis hypothesis regarding the production of intervocalic /d/ will behave in the same way as the retention of intervocalic /d/ under the same internal constraints governing sociolinguistic variation.

variable production of intervocalic /d/. Alternation in the production of the most frequent tokens is high frequency according to the analysis of the tokens *de* 'off/from', *lado* 'side', and these tokens, alternation in the production of these tokens is the dominant tendency.

Group according to the statistical analysis of the data is 42- to 47- month old children. The youngest children favor variable pronunciation of intervocalic /d/. The findings show that alternation is disfavored. Is variability in the production of intervocalic /d/ related to the developmental stage? How can we explain these findings in non-alternating cases? In the analysis of intervocalic /d/ in this same group of children, it was determined that the variable pronunciation is developmental, but sociolinguistic factors influence it from a very early age. The patterns of alternation in the production of intervocalic /d/ are similar. The fact that older children do not show a decrease in the variety of language spoken is evidence that alternation continues to happen and retention of the variable pronunciation is favored in the production of these tokens. In the analysis of the data, it was found that class children begin to retain the variable pronunciation once they are exposed to the variable pronunciation. This is corroborated by the results of the analysis of non-alternating cases: the preponderance of non-alternating cases of intervocalic /d/ (236 tokens out of 250) shows that even though variability in the production of these tokens did show alternation in the production of highly frequent tokens as dis-

groups, dictionary frequency as well as corpus frequency, play a crucial role for predicting variable pronunciation of intervocalic /d/. This piece of evidence will be an important element in the construction of an explanation that we will develop in the next section.

According to the hypothesis presented at the end of section 1, we need to answer whether we find variability in the production of intervocalic /d/ in different age groups. Alternating pronunciation across all age groups could be an indication of acquisition of variable rules, but we will have to explore further pieces of evidence to support this position with solid arguments. Figure 1 shows the patterns of alternation across age groups.

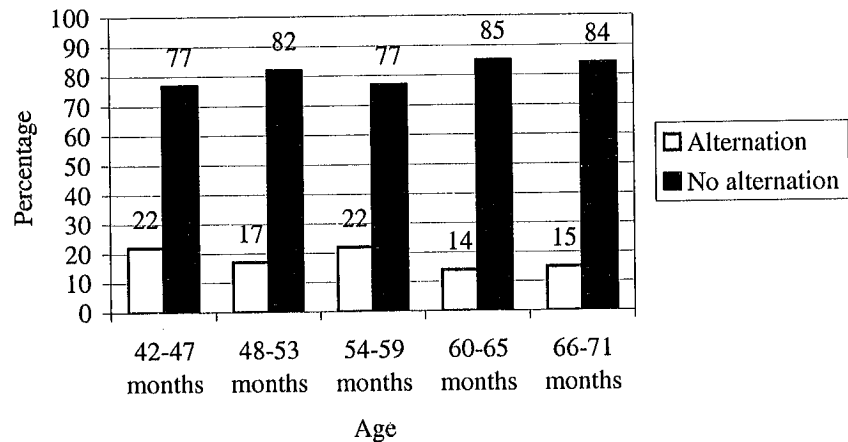


Figure 1. Pattern of alternation in the production of intervocalic /d/ across age groups

Because of the probabilistic analysis presented in section 3.1 we know that 42- to 47- month old children are more likely to produce variable pronunciation of intervocalic /d/. However, figure 1 shows similar percentages across all age groups regarding alternating pronunciation of intervocalic /d/. Recall that alternation is more common in frequent tokens regardless of age group and this fact is reflected in Figure 1. This result could be an indication supporting the variable rule analysis hypothesis because all tokens containing intervocalic /d/ will behave in the same fashion given the internal and external constraints governing sociolinguistic variation of intervocalic /d/ in

is to determine whether children acquire the variable pronunciation in a case-by-case fashion or acquire the variable pronunciation as a result of the hypotheses presented. The results indicate that the two factors

Venezuelan children's speech. However, we need to analyze with more detail the role of frequency in the acquisition of sociolinguistic variables.

Can we show that frequency is playing a role in the spreading of variation in a case-by-case fashion? So far we have seen that the probabilistic analysis has provided good indications that alternation is more likely to affect frequent words. Following this finding, we would have to accept that variability is not having the same effect given the same internal and external constraints. In other words, if children were acquiring variable rules we would expect variability any time the conditions are satisfied, but this is not the result that we are observing in the data examined. Figures 2 and 3 show the pattern of alternation according to dictionary and corpus frequencies, respectively.

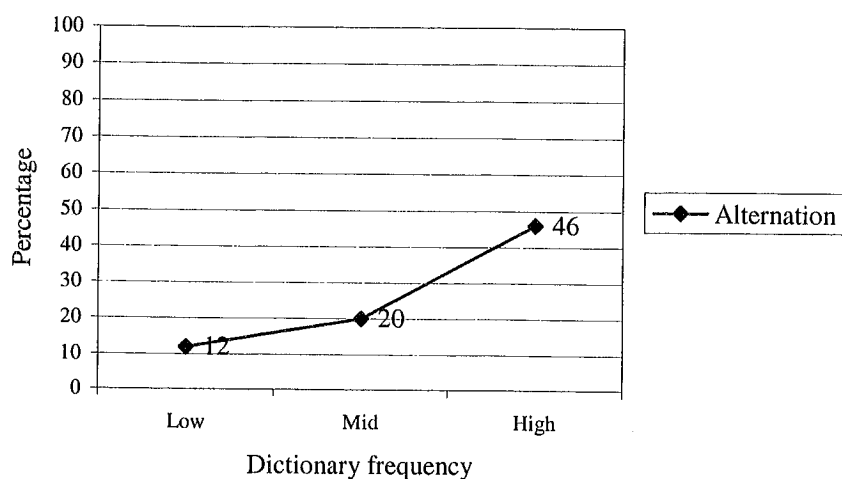


Figure 2. Pattern of alternation according to the factor group dictionary frequency.

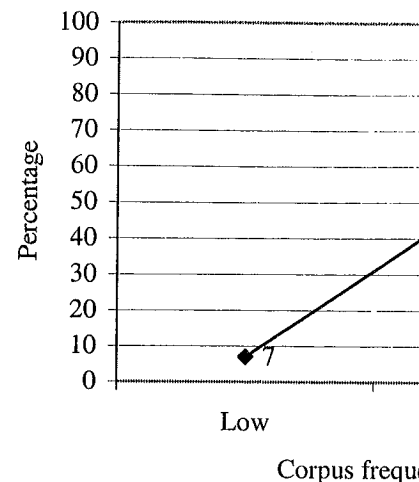


Figure 3. Pattern of alternation according to corpus frequency.

As can be seen in both Figures 2 and 3, the percentage of intervocalic /d/ increases in frequent words. This suggests that sociolinguistic variation is first acquired in frequent words, then it spreads to less frequent ones. This pattern of acquisition of sociolinguistic variables is consistent with the case-by-case rule hypothesis.

In summary, the statistical results show that sociolinguistic variation is first acquired in frequent words. Otherwise, we would have had the same percentage of variation any time the phonetic context and the internal constraints were satisfied. This pattern of acquisition of sociolinguistic variables is consistent with the case-by-case rule hypothesis.

5. Conclusions

The main goal of this investigation was to determine how sociolinguistic variables are acquired by individual learners. To do this we have explored whether the same variables are acquired at the same time across age groups. We also have explored whether the same variables, following Bybee's usage-based model of language change, would spread first in frequent ind-

... need to analyze with more
of sociolinguistic variables.
... role in the spreading of varia-
... ve seen that the probabilistic
... alternation is more likely to
... we would have to accept that
... n the same internal and exter-
... re acquiring variable rules we
... ns are satisfied, but this is not
... amined. Figures 2 and 3 show
... nary and corpus frequencies,

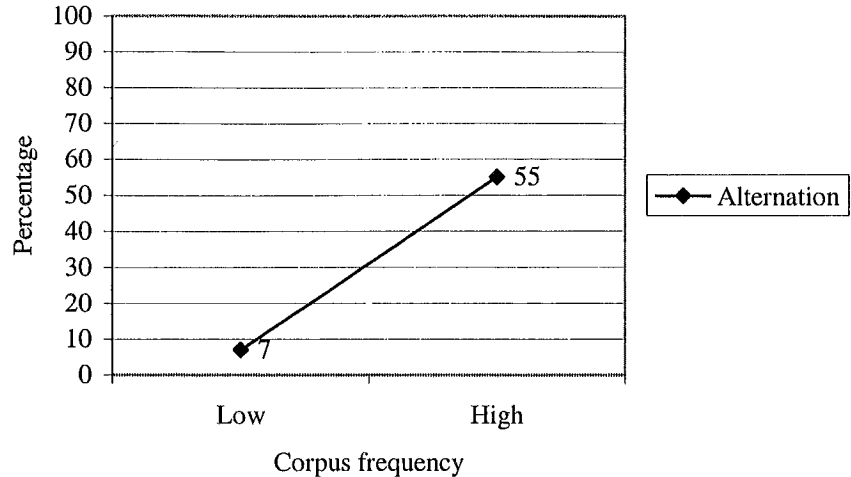


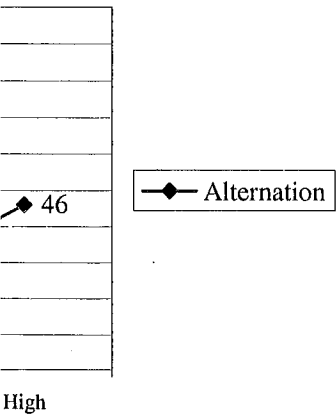
Figure 3. Pattern of alternation according to the factor group corpus frequency.

As can be seen in both Figures 2 and 3, alternating pronunciation of intervocalic /d/ increases in frequent tokens. These findings allow arguing that sociolinguistic variation is first observed in high frequent words and then it spreads to less frequent ones. This would mean that sociolinguistic variables are acquired in a case-by-case fashion contradicting the variable rule hypothesis.

In summary, the statistical results presented here reveal that sociolinguistic variation is first acquired in highly frequent individual lexical forms. Otherwise, we would have had to find alternating pronunciation any time the phonetic context and the internal and external constraints governing variable production of intervocalic /d/ are satisfied.

5. Conclusions

The main goal of this investigation was to answer whether sociolinguistic variables are acquired by individual lexical forms or variable rules. In order to do this we have explored whether the pattern of variation behaves the same across age groups. We also have observed the role of frequency following Bybee's usage-based model of phonology according to which variation would spread first in frequent individual lexical forms. Contrary to our



... the factor group dictionary fre-

expectations, the first question regarding alternating pronunciation across age groups shows that variability is found in all age groups. It is important to point out that variability is more likely to happen in frequent tokens regardless of age. The probabilistic analysis reveals that the most important predictors of alternating pronunciation of intervocalic /d/ are the factor groups dictionary frequency and corpus frequency. This finding is the most relevant piece of evidence of our analysis for supporting the case-by-case hypothesis according to which variability is first found in highly frequent tokens.

In short, the statistical analysis strongly supports the idea that not all the tokens where the conditions of the variable rule are met behave in the same fashion. Frequency effects explain such contradictory behavior providing a crucial argument to support that children acquire sociolinguistic variables in a case-by-case fashion. These results are consistent with Chevrot, Beaud, and Varga's (2000) interactionist approach for explaining acquisition of variable phonology.

Notes

1. This data collection project was supported by the Consejo de Desarrollo Científico y Humanístico of the Universidad Central de Venezuela (grant #: 07-33.3737.96).
2. We also tried to use the *Diccionario de frecuencias de las unidades lingüísticas del castellano* prepared by Alameda y Cuetos (1995), but we found that many common words (e.g., *donde* 'where', *todavía* 'yet', *nada* 'anything', *todas* 'all', etc) are not included in the dictionary creating difficulties for performing a comprehensive analysis.
3. In the results section (see section 4), you will find that this dictionary of frequency based on print sources also reflect very closely frequency in speech. Both dictionary frequency and corpus frequency pattern together in the analysis of acquisition of sociolinguistic variables.
4. The data in Table 1 is organized as follows: the left-hand column shows only the significant factor group in order of selection. The next column indicates the factor values for each factor group selected. The column identified as "No of tokens" has two numbers separated by a slash: the first one refers to the cases of intervocalic /d/ produced variably within the factor considered; the second one indicates the total of tokens of intervocalic /d/ within that same factor or variant. The next column gives the percentage of intervocalic /d/ tokens produced with alternation. The weight column specifies the probabilistic weight for each factor within each group. The range of the weight varies from 0 to 1.

References

- Alameda, José Ramón and Fernando Cuetos
1995 *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Oviedo: Universidad de Oviedo.
- Bybee, Joan
2001 *Phonology and Language*. Cambridge: Cambridge University Press.
- Cedergren, Henrietta
1979 *La elisión de la /d/: un ensayo de la Academia Puertorriqueña de la Lengua Española*.
- Cedergren, Henrietta and David Sankoff
1974 Variable rules: performance and acquisition. *Language* 50: 333-355.
- Chambers, J.
1988 *Acquisition of phonological variables: a quantitative analysis of the acquisition of sociolinguistic variables*. Alan Thomas and Martin Brumfit (eds.), *Linguistic Matters*.
- Chevrot, Jean-Pierre; Laurence Beaud; and Juan Manuel Sosa
2000 Developmental data on French children's acquisition of the consonantal word-final /R/. *Linguistics* 38: 295-319.
- Díaz-Campos, Manuel
2001 *Acquisition of phonological variables: a quantitative analysis of the acquisition of sociolinguistic variables in Venezuelan children's speech*. Juan Manuel Sosa (ed.), *Acquisition of Sociolinguistic Variables*. University of Texas Press.
- D'Introno, Francesco y Juan Manuel Sosa
1986 *Elisión de la /d/ en el español de Venezuela: datos empíricos e implicaciones teóricas*. Juan Manuel Sosa (ed.), *Español del Caribe*, Rafael Marín y Juan Manuel Sosa (eds.), *Actas del Simposio de Sociolingüística*.
- Eckert, Penelope
2000 *Linguistic Variation as a Sociolinguistic Process*. Blackwell.
- Julliard, Alphonse G. and Eugenio Chang-Ross
1964 *Frequency Dictionary of Spanish*. University of Chicago Press.
- Kerswill, Paul
1996 Children, adolescents, and the acquisition of sociolinguistic variables. *Language and Change* 8: 177-202.

References

- Alameda, José Ramón and Fernando Cuetos.
1995 *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Oviedo: Universidad de Oviedo.
- Bybee, Joan
2001 *Phonology and Language Use*. Cambridge: Cambridge University Press.
- Cedergren, Henrietta
1979 La elisión de la /d/: un ensayo de comparación dialectal. *Boletín de la Academia Puertorriqueña de la Lengua Española* 7: 19-29.
- Cedergren, Henrietta and David Sankoff
1974 Variable rules: performance as a statistical reflection of competence. *Language* 50: 333-355.
- Chambers, J.
1988 Acquisition of phonological variants. In *Methods in Dialectology*, Alan Thomas and Martin Ball (eds.), 650-665. Clevedon: Multilingual Matters.
- Chevrot, Jean-Pierre; Laurence Beaud; and Renata Varga
2000 Developmental data on French sociolinguistic variable: Post-consonantal word-final /R/. *Language Variation and Change* 12: 295-319.
- Díaz-Campos, Manuel
2001 Acquisition of phonological structure and sociolinguistic variables: a quantitative analysis of Spanish consonant weakening in Venezuelan children's speech. Ph.D. diss., The Ohio State University.
- D'Introno, Francesco y Juan Manuel Sosa
1986 Elisión de la /d/ en el español de Caracas: Aspectos sociolingüísticos e implicaciones teóricas. In *Estudios sobre la Fonología del Español del Caribe*, Rafael Nuñez Cedeño; Iraset Páez Urdaneta; and Jorge M. Guitart (eds.), 135-163. Caracas: Casa de Bello.
- Eckert, Penelope
2000 *Linguistic Variation as a Social Practice*. Cambridge, Massachusetts: Blackwell.
- Jullian, Alphonse G. and Eugenio Chang-Rodríguez
1964 *Frequency Dictionary of Spanish Words*. The Hague: Mouton.
- Kerswill, Paul
1996 Children, adolescents, and language change. *Language Variation and Change* 8: 177-202.

- Kovac, Ceil and H. D. Adamson
 1981 Variation theory and first language acquisition. In *Variation Omnibus*, David Sankoff and Henrietta Cedergren (eds.), 403-410. Edmonton, Alberta: Linguistic Research, Inc.
- Labov, William
 1964 Stages in the acquisition of standard English. In *Social Dialects and Language Learning*, Roger Shuy, Alva Davis, and Robert Hogan (eds.), 77-104. Champaign: National Council of Teachers of English.
 1972 *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
 1989 The child as linguistic historian. *Language, Variation, and Change* 1: 85-97.
- Patterson, Janet
 1993 The development of sociolinguistic phonological variation patterns for (ING) in young children. *Dissertation Abstracts International* 53:3, 3196A.
- Payne, Arvilla
 1980 Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In *Locating Language in Time and Space*, William Labov (ed.), 143-178. New York: Academic Press.
- Roberts, Julie
 1994 Acquisition of variable rules: (-t,d) deletion and (ing) production in preschool children. Ph.D. diss., University of Pennsylvania.
 1997a Hitting a moving target: Acquisition of sound change in progress by Philadelphia children. *Language, Variation, and Change* 9: 249-266.
 1997b Acquisition of variable rules: A study of (-t,d) deletion in preschool children. *Journal of Child Language* 24: 351-372.
- Roberts, Julie and William Labov
 1995 Learning to talk Philadelphian: Acquisition of short a by preschool children. *Language, Variation, and Change* 7: 101-112.

Interplay between phonetic constraints in the degree of stops: Comparing intervocalic /g/ in Spanish and English

Marta Ortega-Llebaria

1. Introduction

Explaining speech variability is a long-standing problem in linguistics that has been addressed from different perspectives. The great variability encountered in the degree of voiced stops in Spanish (Quilis 1963) has often been treated in a categorical manner, reduced to two categories, spirantized and lenited. Their contextual alternations have been explained either by the underlying realization of the stop as a spirant (1989; Mascaró 1991), or the underlying realization of the stop as a lenited stop by assimilating the [-continuant] feature of the preceding consonant (Lozano 1979: 14-29).

However, recent studies have taken into account the degree of lenition and have described the phenomenon whose variability is conditioned by the principles of gestural reorganization. For example, Quilis (1998) found that the degree of spirantization varied according to stress and vowel context. In Spanish, spirantized in unstressed syllables flanked by vowels, e.g., /súgus/ 'candy', and least spirantized in stressed syllables, e.g., /amagár/ 'to show, to inform'. After finding de-occlusivization of /m/ in Spanish, Quilis (1998) proposed that spirantization of intervocalic voiced stops and de-occlusivization are patterns that obey the same gestural reorganization in which the gestural magnitude varies in intervocalic contexts.