Effects of variable input in the acquisition of plural in two dialects of Spanish

Karen Miller\textsuperscript{a}, Cristina Schmitt\textsuperscript{b,*}

\textsuperscript{a}Calvin College, Department of Spanish, Hiemenga Hall, Grand Rapids, MI, 49506 United States
\textsuperscript{b}Michigan State University, Department of Linguistics, Germanic, Slavic, Asian and African Languages, East Lansing, MI 48824, United States

Received 20 September 2007; received in revised form 30 May 2008; accepted 7 April 2009
Available online 22 July 2009

Abstract

In Mexico City Spanish, plural morphology is always overtly realized but in Chilean Spanish, due to a lenition process that targets syllable-final [s], plural morphology is not always realized, and the conditions of overt realization of plural morphology are subject to both linguistic and extra-linguistic factors. In this paper we show that variability in the input for grammatical morphology causes the performance of Chilean and Mexican children to differ from each other and from the adult control groups in production and comprehension tasks in ways that correlate with the reliability of plural marking in the input. Our results support Yang (2002)’s proposal that variability in the input does cause some children to assume a different grammar from the adult.

\textcopyright{} 2009 Elsevier B.V. All rights reserved.

Keywords: Plural; Spanish; Acquisition; Variable input; Comprehension

1. Introduction

It is a truism that the input that a child receives is crucial to determine what language the child acquires. We are not surprised that a child exposed to English learns English, and a child exposed to Spanish learns Spanish. However, beyond this simple characterization of the relation between input and acquisition lies an important empirical question: exactly how does the input determine the grammar that is acquired by the child? We know that the linguistic representations the child constructs are more abstract than simply the surface distributions of linguistic forms. We also know that what the child hears may not transparently and categorically reflect the adult grammar(s). What we do not know is to what extent differences in the input affect children’s paths to their adult grammatical representations.

\textsuperscript{☆} This study was funded by NSF Doctoral Dissertation Improvement Grant #0446769, NSF Grant BCS-0126502 and the Michigan State University Graduate Student Research Enhancement Award. We thank the following schools in Punta Arenas, Chile: Colegio Aleman, Colegio Britanico, Colegio Pierre Faure, Junta Nacional de Jardines Infantiles de Chile (JUNJI), Jardin Bambi, Jardin Las Charitas, Escuela 18 de Septiembre, Colegio Miguel de Cervantes, y Universidad de Magallanes and also in Mexico: Centro de Desarrollo Infantil (CENDI) and the Universidad Autonoma Metropolitana (UAM), Campus Iztapalapa of Mexico, D.F. We especially thank John Grinstead and Antoinette Hawayek for their assistance in Mexico City and for their comments and suggestions on this work. Thanks to the following research assistants: Rodrigo Cardenas, Cynthia Corona, Marena Garcia, Katerina French, Edgardo Mansilla, Erika Mendoza, Andrew Sanford, Heriberto Sierra, and Pascale Schnitzer. Finally, we thank Ana Perez-Leroux, Alan Munn and the members of the Michigan State University Language Acquisition Laboratory and the audience at FAVS, York 2007.

* Corresponding author.

\textit{E-mail address:} schmit12@msu.edu (C. Schmitt).

0024-3841/S – see front matter \textcopyright{} 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.lingua.2008.05.009
In this paper we ask how unreliable and ambiguous input affects the acquisition of grammatical morphology. The empirical domain is the acquisition of plural morphology in Spanish, more specifically Chilean Spanish and Mexico City Spanish (henceforth Mexican Spanish). These dialects differ in the overt realization of plural morphology. In Mexican Spanish, plural morphology is overtly marked as [s] on nouns, determiners and adjectives within the noun phrase (Lipski, 1999; Moreno de Alba, 1994), as illustrated in (1a); singular morphology, on the other hand, is not overtly marked, as illustrated in (1b). In Chilean Spanish, on the other hand, there is a substantial amount of socio-linguistic variation in the production of plural morphology, due to a well-known lenition process (common in many dialects of Spanish) that affects all syllable-final /s/, reducing [s] to a simple aspiration [h] and sometimes to zero (Ø) (Cepeda, 1995; Fox, 2006; Miller, 2007), as illustrated in (2a). As in Mexican Spanish, singular is not overtly marked (2b).

(1) a. unas niñas pequeñas
   ‘some small girls’

   una niña pequeña
   ‘a small girl’

(2) a. una[s, h, Ø] niña[s, h, Ø] pequeña[s, h, Ø]
   ‘some small girls’

   una niña pequeña
   ‘a small girl’

Lenition of syllable-final /s/ masks the overt realization of plural morphology in the adult language creating a fair amount of non-categorical ambiguity between singular and plural noun phrases, especially in the indefinite feminine (1) and the definite feminine forms where the singular and the plural form differ only by the overt marking of plural (la ‘the–fem–sg’ vs. las ‘the–fem–pl’). Although all Chilean children are exposed to some overt plural marking, some children receive input with more overt plural marking than others. Plural contexts in which the plural morphology is overtly realized provide evidence for overt number marking. Conversely, contexts in which plural morphology is omitted due to lenition provide evidence that there is no overt marking of plural or singular morphology in the language.

The difference between the two varieties of Spanish is both quantitative (Chilean children are exposed to less overt plural marking) and qualitative. Not only are Chilean children exposed to less overt realization of plural marking, but the overt realization is not categorical and, arguably, less reliable. Overt plural morphology can be absent in the noun phrase, marked only on the determiner, the noun or neither. Variability (within and across speakers) adds ambiguity to the input the child is exposed to, making some input unreliable for a particular grammatical generalization although perfectly reliable for learning some other property of the language (information about socio-linguistic variation, for example).

Although much important work in acquisition has addressed the problem of ambiguity in the input in a homogeneous speech community (Fodor, 2001; Gibson and Wexler, 1994; among many others), very few generative linguists have dealt with probabilistic properties of natural languages, except when dealing with language change (Kroch, 1994, 2000; Pintzuk, 2002). Recently, however there has been a new and welcome impulse towards a formal treatment of synchronic variation in generative linguistics (Adger and Smith, 2005; Parrott, 2006; Tortora, 2006; Adger, 2006; Roeper, 1999, 2007; Green and Roeper, 2008, papers in this issue). This move has allowed acquisition researchers to relax the idealization that the child is exposed to an unvarying, homogeneous speech community and provides theoretical tools with which to examine more carefully non-categorical properties of the input and their impact on acquisition. After all, as pointed out by Wilson and Henry (1998), variability in the input is actually the norm, and the acquisition device must be able to cope with it.

In order to relax the homogeneous speech community idealization and to accommodate the non-categorical properties of the input with the categorical properties of grammars, we need a model that can take both parametric and statistical properties into consideration. The Variational Model (Yang, 1999, 2002) is exactly such a model. Building on the idea that the learner makes hypotheses within the limits of UG that are punished or rewarded depending on their ability to account for properties of the input, Yang adds a component to the learning model that is sensitive to statistical properties of the input. This sensitivity to properties of the input can explain why certain parameters are set much
earlier than others. If the input is unambiguous and frequent, acquisition happens early because a particular value of a parameter is consistently rewarded. If the input is ambiguous, the child may take longer to set a parameter; or may set a parameter incorrectly; or may allow for two opposite values of a parameter to coexist. Yang’s model allows for a learner to maintain two opposite values or choose the non-adult grammar value, making room for variation and/or change. In other words, the model allows us to predict that the more reliable (categorical and unambiguous) the input, the faster the child will converge on a grammar that is identical to the adult grammar. Conversely, the more ambiguous the input is, the longer children will take to converge on the target grammar if he/she converges at all.

Attempts to examine the role of statistical properties of the input in language acquisition are not new. In the acquisition of morphology and syntax, there is a long tradition of examining the effects of statistical properties (Brown, 1973; Gathercole, 1986; Demuth, 1989, 2007; Stromswold, 1995; among others) as they correlate with the development of syntactic and morphological properties in different languages or the development of different properties within the same language. Very few studies, however, have examined the effect of variable and unreliable input on language acquisition of closely related dialects (Roberts, 1994; Demuth and Ellis, 2008; Hudson Kam and Newport, 2005; Johnson, 2005; Johnson et al., 2005; Smith et al., 2006; Smith, 2007; Miller and Schmitt, 2004, 2006a,b, 2009). Within this group, even fewer studies have dealt with the consequences of unreliable input for comprehension (Johnson, 2005; Miller and Schmitt, 2004, 2006a,b, 2009; Miller, 2007). In this sort of study, the effect of both the properties of the input and the distributional frequency of those properties must be examined.

With respect to number morphology, since the work by Brown (1973), plural morphology is considered part of the child’s early system. English-speaking children by age 2 can produce plural morphology 90% of the time. For Spanish, researchers report that consistent use is found also around age 2 (Marrero and Aguierre, 2003). The work by Berko (1958), later replicated both in English and other languages by younger children (Kernan and Blount, 1966; Anisfeld and Tucker, 1967; Perez-Pereira, 1989; Ferenz and Prasada, 2002; Cantú-Sánchez and Grinstead, 2004 for Spanish), showed also that children by 4 years of age had acquired plural morphology and its less frequent allomorphs. In other words, children can produce all the allomorphs of the plural morpheme at adult levels by 4 years of age in English and Spanish when the input is reliable.

Interestingly, there is much less research on the development of plural morpheme comprehension (Barner et al., 2007). For standard English, the empirical evidence shows that by 24 months children can use nominal and verbal number morphology together to distinguish ‘one’ vs. ‘more than one’ and by 36 months number information on the noun phrase alone enables children to distinguish ‘one’ vs. ‘more than one’ in novel words. For Mexican Spanish (where the input is reliable) we have found that at least by 3.5 years of age, children can use nominal plural morphology to distinguish ‘one’ vs. ‘more than one’ (Miller, 2007; Munn et al., 2006). These results suggest that in languages like Mexican Spanish and English, where the input for plural morphology is reliable, the overall frequency of the plural morpheme (e.g., the number of tokens of plural marking) does not determine age of acquisition. Although in Spanish there are more tokens of the plural morpheme because it occurs on every element inside the noun phrase, there does not seem to be a significant age difference in the acquisition of plural morphology in both languages. In other words, overall frequency of the realization of plural morphology is not a factor that delays acquisition.1

In this paper, we compare Chilean children to Mexican children in order to test Yang’s Variational approach and we ask whether variability in the input for plural morphology causes the performance of Chilean and Mexican children to differ from each other and from the adult control groups in ways that correlate with the reliability of plural marking in the input. Our results support Yang’s proposal that variability in the input does cause some children to assume a different grammar from the adult and also support the hypothesis that ambiguity in the input delays the convergence into the target language.

The paper is divided as follows. Section 2 provides an introduction to the problem of morphology acquisition, a brief description of the two dialects, and presents our hypotheses and predictions. Section 3 describes production results from Miller (2007). Section 4 presents two experiments on the acquisition of number and section 5 discusses the results in light of the hypotheses and predictions.

---

1 A reviewer notes that one could argue that the plural morpheme is more frequent in Spanish but is also more complex since it is also a mark of agreement, and this could perhaps explain the lack of apparent differences in the acquisition of plural morphology in English and Spanish. Although this is an interesting possibility and one that highlights the complexity of input properties that must be taken into account, we find this possibility unlikely because first language acquisition studies of agreement morphology in Spanish speaking children show mastery of agreement very early and also very few errors (see Brandani and Álvarez, 2004; Álvarez and Brandani, 2005; Cantú-Sánchez and Grinstead, 2004, among others). Although these are not large studies they suggest that age of acquisition when the input is reliable is very similar in both Spanish and English.
2. Background

2.1. The problem of morphological acquisition

To know a piece of grammatical morphology is to know how a form maps to a syntactic feature, and how the syntactic feature is mapped into a semantics and, therefore, an interpretation. In the simplest case, one form maps to one syntactic feature and the syntactic feature to a unique semantics, and the acquisition problem is reduced to figuring out the feature with which a particular form is associated. Assuming the set of features to be universal, children have to learn which feature is being mapped to which form, and what semantics is associated to it. In most work on the acquisition of functional morphology, researchers have assumed that production at adult levels is a very clear test for mastery of inflectional morphology. Such a view assumes that the mapping from syntax to semantics is transparent and/or comes for free in the acquisition process.

However, a particular form is not always associated to the exact same features in all contexts. The one-to-one mapping between a form and a syntactic feature does not always obtain. Assuming Distributed Morphology (DM) ([Halle and Marantz, 1994](#)) the one-to-many mapping becomes quite clear. In DM, morphological forms are inserted late in the derivation. The Subset Condition for Vocabulary Insertion ([Halle, 1997](#), among others) just says that a vocabulary item cannot have more features than the position it is inserted in, but it could have less, if there is not a better candidate for insertion. What this means is that the same form is not always the spell-out of the same set of features. If we only consider spontaneous production, it may be that children use a particular morphological piece in just a subset of the cases the adults would use it and his/her use in those contexts can reach adult levels.

Furthermore, syntactic features can be interpretable or uninterpretable in the semantics. If interpretable, a syntactic feature may or may not be directly associated with a particular semantics. If uninterpretable, a feature is not visible to the semantic component directly, but since all uninterpretable features are valued by an interpretable feature (with perhaps the exception of Case) in an Agree relation, uninterpretable features are indirectly associated to an interpretation. The interpretation can come from being associated to a well-defined lexical semantics. However, this is not always the case. In many cases, the interpretation of a particular syntactic feature (or its absence) arises by implicature. Schematically, the idea is that, if we have a pair of features (X and Y) where X is marked and Y is unmarked, if Y is being used, it must be the case that the conditions for X have not been satisfied. By implicature, we can interpret Y as not X.

Turning to the case of plurality, predictions about children’s behavior will depend on the hypothesis we make about the adult system. A system where there is a one-to-one correspondence between a semantic singular and a syntactic singular and a semantic plural and a syntactic plural is too simplistic if we consider mass nouns, pluralia tantum cases and agreement markers. Even ignoring mass nouns, pluralia tantum and the issue of agreement, the simplistic view is problematic.

Consider the configuration in (3).

(3)  
```
      NumP
     /     
    [-SG]  nP
     /      
    n      ROOT
     /     
   [COUNT]
```

(4) a. substitute [-s] [-SG] for [s] is associated to syntactic feature [-SG]  
b. no vocabulary item for [count] or for [+SG]

(5) a. substitute [-s] [COUNT] for [s] is associated to syntactic feature [COUNT]  
b. no vocabulary item for [+SG] or for [-SG]

---

2 For concreteness we can assume Adger (2006) who defines the agreement relation in the following way: an agreement-chain is a pair of lexical items, where the uninterpretable features of one LI are a subset of the interpretable features of the other.
In (3) we assume that n is the locus for both gender and count/mass distinctions (see Saab, 2004; Doetjes, 2007, and references therein). The presence of the syntactic feature [–C0SG] in (3) co-occurs with the feature [COUNT]. The question is what [–s] stands for. In this structure, /–s/ can easily be interpreted by the child as the realization of either [–SG] or [COUNT] as in (4) or (5). Either [–s] morpheme could be inserted in (3). In other words, the child could be treating the morphological realization of plural simply as an overt indication of a count noun with the distinction between singular and plural made by the element in the number head. After all, in most cases, grammatical number is redundant information because the D will guarantee an atomic or a non-atomic interpretation.

On the mapping between the syntax to the semantics there are also complications, depending on our theory of plurality. For example, if we treat the syntactic features singular and plural as being associated to two independent semantic features, we may predict better performance in comprehension tasks than if we assume a privative system in which the interpretation of one of these syntactic features arises through implicature, arguably harder for children (Sauerland et al., 2005; Reinhart, 1999; Noveck, 2001; among others).

In this paper, we assume that all instances of overt plural marking on the noun phrase in both dialects of Spanish are instances of the morpho-phonological realization of an uninterpretable feature that needs to be valued by the interpretable number feature in a functional node above nP. The count/mass distinction is made at the level of nP in the adult grammar. Second, uninterpretable features that have been valued are visible at the PF interface and may be used in comprehension tasks to recover the relevant interpretable feature. As far as the semantic value of interpretable plural and singular, for concreteness, we assume a two-way distinction between [+SG, –C0SG] where the positive value for singular is associated to the semantics in (6) and the negative value is associated to its negation, as in Harbour (2007).

\[\lambda P\alpha \chi [P(\chi) \land \text{atom } (\chi)]\]

For the purpose of this paper it suffices to say that Number morphology is interpreted once in the DP. The feature [+SG] will be associated with atomic interpretations and [–SG] will be associated with non-atomic interpretations of count nouns.

2.2. Properties of the input

While in Mexican Spanish plural morphology is always present in production, the data for dialects of Spanish are instances of the morpho-phonological realization of an uninterpretable feature that needs to be valued by the interpretable number feature in a functional node above nP. The count/mass distinction is made at the level of nP in the adult grammar. Second, uninterpretable features that have been valued are visible at the PF interface and may be used in comprehension tasks to recover the relevant interpretable feature. As far as the semantic value of interpretable plural and singular, for concreteness, we assume a two-way distinction between [+SG, –SG] where the positive value for singular is associated to the semantics in (6) and the negative value is associated to its negation, as in Harbour (2007).

\[\lambda P\alpha \chi [P(\chi) \land \text{atom } (\chi)]\]

For the purpose of this paper it suffices to say that Number morphology is interpreted once in the DP. The feature [+SG] will be associated with atomic interpretations and [–SG] will be associated with non-atomic interpretations of count nouns.

2.2. Properties of the input

While in Mexican Spanish plural morphology is always present in production, the data for dialects of Spanish with lenition is a lot more complex (Lipski, 1999). For Chilean Spanish, Cepeda (1995) reported that the highest rate of syllable-final /s/ omission, due to the lenition process, occurs with the plural marker /–s/, rather than with non-morphological /s/. She found that for 34 adult Chilean speakers, plural /–s/ was omitted on 54% of all nouns (4580 tokens) and on 60% of all post-nominal adjectives (785 tokens). Omission of plural marking in determiners (e.g. *esos ‘these.pl’, unos ‘some.pl’, las ‘the.f.pl’) was lower than in nouns and adjectives. Within determiners, omissions were lower (13%) in monosyllabic determiners (e.g. las, ‘the.fem.pl’, mis ‘my.pl’) than in polysyllabic determiners (e.g. unos, ‘some.pl’ algunos ‘some.pl’), which occurred 34% of the time. Moreover, omission rates were linked to social class, with omissions being more likely in the language of working-class speakers compared to middle-class and upper-class speakers, as illustrated in Table 1 (adapted from Cepeda, 1995).3

Lenition of /–s/ introduces a certain degree of unreliability into the input that does not exist in either English or Mexican Spanish with respect to the mapping between forms and the feature [+/–SG]. There is no consistent and systematic evidence for and against overt marking of the plural/singular distinction or agreement, and the relation

3 Cepeda (1995) does not separate [s] from [h]. Miller (2007) shows that the majority of plurals are realized as [h] and not [s], but [s] is very frequent in bare plurals.
between [s] and [h] and the feature [−sg] is sparser, since a considerable number of nouns that are interpreted as plurals are not morphologically different from the ones interpreted as singular. For example, it is common for expressions such as (7a) to appear as (7b). In this case the definite and the noun are identical to the singular form in (7c).

(7)  a. todas las manañas
    all-FEM-PL the-FEM-PL morning-FEM-PL
    all the mornings
  b. toda la mañana
    all the mornings
  c. toda la mañana
    all morning long

2.3. The acquisition task and predictions

Based on the facts described above, we have been conducting a series of experiments using Chilean and Mexican children to test the effects of the input variation both in comprehension and production (Miller and Schmitt, 2004, 2006a,b, 2009; Miller, 2007). The overall logic of our experiments is the following: we test Chilean children of two different social classes (who receive different amounts of input variation) and compare them with Mexican children (whose input is not variable). The Mexican children therefore serve as a control group with respect to the amount of input variation.

Under the assumptions made in the previous section, for both the Mexican and the Chilean child, the first task is to determine whether their language makes a grammatical distinction between [+sg] and [−sg] forms or not. Once that is settled, the task is the same: to determine how the mapping between syntax and semantics and the mapping from form to syntax get done. We can use the Variational Model to make two types of predictions: we can make predictions for different groups and also for individual children within groups.

For the Mexican child (and in fact for the English child), we believe that some form of syntactic bootstrapping is at work in the acquisition of the number morphology. The grammar that makes a plural/singular distinction will be systematically rewarded in these languages by the categorical presence of plural morphology in noun phrases that have determiners that impose a particular non-atomic restriction on the set (many, some, two, three, etc.). The absence of plural morphology, on the other hand, is obligatory when the determiner restricts the set to the atoms (the indefinite singular ‘a’ in English form example), as illustrated in (8) where the starred examples would provide evidence against overt plural marking and are not possible in Mexican Spanish.

(8)  a. dos galletas/*dos galleta
    two cookie-PL/*two cookie
  b. una galleta/*una galletas
    a cookie-SG/*a cookie-PL

Given the nature of the input we expect that individually and as a group they will perform well in both production and comprehension tasks, since their grammar should be the grammar that makes an overt distinction between singular and plural.

Due to the lenition process, the Chilean situation can lead to a problem of mapping forms to [+sg] and [−sg]. For example, noun phrases that appear in DPs that restrict the set to non-atomic properties and in DPs that restrict the set to atomic properties will have the same form, as illustrated below:

(9)   dos galleta
    two cookie
    una galleta
    a cookie

In other words, the child cannot reliably use a bootstrapping strategy to acquire the overt marking of the singular/plural distinction. This can lead the child to entertain, for a much longer period, the competition
between a grammar with a grammaticalized plural/singular distinction and a grammar without a grammaticalized singular–plural distinction, since the same determiner will appear some of the time with overt plural morphology in the noun phrase and some of the time without plural morphology in the noun phrase, rewarding both grammars.

Given that in the Chilean case the input has a variable amount of ambiguity, the Variational Model allows three possible outcomes:

(i) **Positive regularization:** the Chilean child may assume that the language makes the distinction and marks it overtly.

(ii) **Negative regularization:** the Chilean child may assume that the language does not make the distinction overtly.

(iii) **Grammar competition:** the child will assume two grammars: in some cases that language makes the distinction and in some cases that the language does not make the distinction. The choice of grammar will depend on the weight a particular grammar has at a particular point and that depends on how frequently it is rewarded.

It is important to note that the three outcomes proposed above are outcomes that speak to the learner’s competence, and not his/her performance. The outcomes predicted by the Variational Model could be thought of as three different knowledge states. Each outcome in turn makes different behavioral predictions for production tasks and comprehension tasks in which the child has to act-out instructions involving singular and plural noun phrases and need to make decisions between singleton sets or plural sets.

If Chilean children regularize the input in a positive way, they will produce more plural morphology than the adults (as attested by Newport, 1999; Hudson Kam and Newport, 2005). In comprehension tasks they will also perform like the Mexican children. We can expect that all children will perform in the same way as Mexican children in comprehension tasks. Whatever problem they may have in comprehension should be identical in all dialects.

If Chilean children have regularized the input in a negative way, they will not produce plural morphology as plural morphology and will not use it in comprehension tasks. Performance in comprehension tasks will be systematically non-target like, and children will rely on other sources of information in forced-task choices, namely linguistic and extra-linguistic context. Here we should find Chilean children behaving very differently from Mexican children. If we find children producing what in the adult language is associated to plural morphology but unable to use it in comprehension tasks, one potential explanation would be that they are using the so-called plural morphology as a marker of something else. A reasonable possibility is that they interpret it as realizing the feature [+COUNT].

If Chilean children have two grammars, performance in production and comprehension tasks could in principle vary for a single individual. The rate of plural morphology produced will vary depending on the input children are exposed to. Based on Cepeda’s findings, we expect ChMC children to produce more plural morphology than ChWC children since the grammar that posits a grammatical distinction between singular/plural will be rewarded more often. With respect to comprehension, again we may find a variable behavior or we may find children systematically using one or the other grammar within a single experiment. In other words, if the child has two grammars, they can pick anyone of the two grammars to parse a particular string. If they choose the one that does not mark grammatical number overtly, they will ignore the plural morphology (or treat it as associated to some other feature). As we discussed above, the so-called plural form may be used purely to unambiguously associate a count feature to the noun phrase. If however they choose the grammar that marks overtly the distinction between singular and plural, they should parse plural morphology as [−SG]. At least across experiments we should find children sometimes ignoring plural morphology and sometimes interpreting it.

These three possible outcomes are dependent on the frequency of the input. Too little plural morphology will lead the child to assume (ii); enough plural morphology will force the child to assume (i) and if the input presents a lot of ambiguity, the model predicts that the child will entertain for a long time two grammars with equal or different

---
4 In this paper, we only test children’s ability to use [+PL] as a associated to a non-atomic set. See Miller and Schmitt (in prep.) for Chilean and Mexican children’s ability to use plural morphology to make a count/mass distinction.
weights, depending again on how much each of the grammars was rewarded.\(^5\) The predictions for groups are also clear. The prediction from the Variational model is that ChWC should have mostly assumed (ii) or (iii), followed by ChMC children and Mexican Children, which should be at ceiling, given the previous studies on production and comprehension of plural morphology.

3. Properties of children’s production

In this section we summarize Miller’s (2007) results which compared overall adult production with the production from 4 to 5 year olds in three different production tasks: a free speech task, a naming task (subjects named plural and singular sets of toys after the prompt “What is here?” (see Miller, 2007 for more details of the experimental design)) and a repetition task (subjects repeated 16 sentences with plural and singular noun phrases in subject and object position).

3.1. Subjects

16 Chilean Working Class (ChWC) children (4.5–5.11: mean 5.2), 8 Chilean Middle Class (ChMC) children (4.8–5.11: mean 5.3), 3 ChWC mothers, 4 ChMC mothers, 8 Mexican Working Class (MexWC) children (4.8–5.9: mean 5.4), 2 MexWC parents and 2 Mexican Middle Class (MexMC) adults (the MexMC adults were parents but not parents of the Mexican children who participated in the study) participated in the study. There are more ChWC children than ChMC children or MexWC children because not all ChWC children were asked to participate in the Free Speech task.

3.2. Procedure

Subjects were tested in their school or in their own homes by a native Spanish-speaking research assistant who was from the same local area as the subjects. During the Free Speech task, the native-speaking research assistant used colloquial forms of the plural morpheme in their own speech. In other words, the Chilean Spanish research assistants used mainly [s], [h] and Ø while the Mexican research assistants used [s]. However, during the Repetition Task, both the Chilean and Mexican research assistants consistently pronounced the plural morpheme as [s]. Given that in Chilean Spanish [h] and Ø are the results of syllable-final /s/ lenition, having the researchers pronounce the plural as [s] provided subjects with the acoustically most salient variant (Fox, 2006). The purpose was to elicit [s] from both child and adult subjects. The Free Speech Task was administered first, followed by a Naming Task and then a Repetition Task. The Repetition Task was carried out last so as not to prime the use of the plural or any of its variant forms in subsequent tasks.

3.3. Results

A total of 1739 plural tokens were collected from the Chilean subjects and a total of 1308 plural tokens were collected from the Mexican subjects. It is first important to note that with respect to the Chilean data, the data was transcribed and then two native-speakers coded independently the plural noun phrases for [s], [h], Ø. Inter-rater reliability between the two Chilean native-speaking coders reached approximately 83.5% of all plural tokens. In other words, the coders disagreed on 285 of the 1739 tokens (16.4%) and disagreement was between a variant of [h] vs. Ø, where one native speaker coded as Ø and the other has [h]. In cases of disagreement between coders (approximately 16.4% of the tokens), the token was coded as Ø under the reasoning that if native-speaking coders could not decide

\(^{5}\) The key issue is to determine what counts as enough and what counts as too little plural morphology. The exact numbers are an empirical issue. One could perhaps sharpen the outcome predictions above by examining the few previous studies that investigate acquisition dependent on less reliable input. As pointed out by a reviewer, Singleton and Newport (2004)’s study of Simon, the deaf child that learned ASL from their non-native ASL parents, is a potential candidate. This study showed that Simon regularized the input when the use by parents was above 70% and did not regularize it when it was below 50%. Although it is tempting to use those numbers to make a more concrete proposal for the acquisition of plural morphology in Spanish, we will refrain from hypothesizing a percentage or a weight necessary and/or sufficient based on these previous studies because it is not clear that children require the same frequencies and/or weights to set parameters for different grammatical properties (see Roeper and Green, this volume).
between zero and a variant of [h], native-speaking children would also have difficulty distinguishing between presence or absence of the plural marker.

Table 2 summarizes the production results: first we see that the MexWC children and adults produce plural morphology nearly 100% of the time in the three tasks: free speech production, repetition and naming task. The Chilean data replicates Cepeda’s 1995 overall findings: ChWC adults produced less plurals than ChMC adults. Chilean children differed from the adults in the same social group in two ways: (i) ChMC children, but not ChWC children, produced significantly more plural morphology than ChMC adults; and (ii) ChWC children produced significantly less plural morphology than ChWC adults in repetition and naming but not in free speech.

These differences must be further investigated since they may result from adult-directed speech (ADS) being different from child-directed speech (CDS) (see Smith et al., 2006 for results along this line) and parents may have used a different register with the investigators. Overall it is very clear that there are more omissions in the speech of ChWC adults than in the speech of ChMC adults. Importantly all adults produced a certain amount of plurals and all children (with exception of 2) produced plural morphology at least once. The results also show that, while ChMC children are sensitive to style (repetition and naming vs. free speech), ChWC children are not, as shown in Fig. 1 below.

If the adult free speech data omission rate is a reliable indicator of children’s grammars, we can expect that both grammars would be equally weighted for the ChMC children but not for the ChWC children, given the Chilean adult numbers.

If children’s production is a better indicator of their knowledge state, the picture goes in the same direction but the difference between the two groups of children will be much wider, since ChWC children seem to omit plural morphology more than 50% and do not seem to be sensitive to register. If sensitivity to register can be used as evidence for grammars in competition, ChMC would arguably have two grammars. The fact that almost all children produced some amount of plural morphology is consistent with the hypothesis that there are two grammars

<table>
<thead>
<tr>
<th>Subjects</th>
<th>[s] or [h] %</th>
<th>N</th>
<th>Zero %</th>
<th>N</th>
<th>N of zero [h], [s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MexMC adults</td>
<td>99</td>
<td>335</td>
<td>1</td>
<td>5</td>
<td>340</td>
</tr>
<tr>
<td>MexWC adults</td>
<td>99</td>
<td>430</td>
<td>1</td>
<td>3</td>
<td>433</td>
</tr>
<tr>
<td>MexWC children</td>
<td>99</td>
<td>532</td>
<td>1</td>
<td>3</td>
<td>535</td>
</tr>
<tr>
<td>ChMC adults</td>
<td>67</td>
<td>433</td>
<td>33</td>
<td>213</td>
<td>646</td>
</tr>
<tr>
<td>ChMC children</td>
<td>83</td>
<td>278</td>
<td>17</td>
<td>58</td>
<td>336</td>
</tr>
<tr>
<td>ChWC adults</td>
<td>56</td>
<td>145</td>
<td>44</td>
<td>113</td>
<td>258</td>
</tr>
<tr>
<td>ChWC children</td>
<td>42</td>
<td>212</td>
<td>58</td>
<td>287</td>
<td>499</td>
</tr>
</tbody>
</table>

Fig. 1. Percentage of plural omission in free speech, repetition and naming tasks.
competing for most children. Furthermore, the different rates of plural omission in the different groups is compatible with the different weights associated to the two grammars. We will call this the Grammars in Competition Hypothesis.

The problem is that given that we are examining performance in production, we cannot guarantee that the source of omissions is necessarily linked to a grammar where the plural/singular distinction is not made. It could be that the amount of plural morphology they hear in the adult input is sufficient to drive children to choose a grammar with plural morphology. The differences between children and adults should perhaps not be linked to a difference in grammars but could be due to performance problems. We will call this the Production Problem Hypothesis. Although further research is required to test this hypothesis, it makes very clear predictions about their behavior in comprehension tasks. If they know all there is to know about plural morphology, then they should do well in comprehension tasks. In order to tease apart which of these two hypotheses has merit, in the next section we report two experiments examining children’s performance in comprehension tasks involving the indefinite determiner.

Under the Grammars in Competition Hypothesis, we expect that overall ChMC children will perform better than ChWC children. Under the Production Problem Hypothesis we expect that ChMC and ChWC children should perform alike.

4. Comprehension of un vs. unos

In order to test comprehension of singular and plural morphology with indefinites we used an act-out task. Consider a situation in which there is a box and some marbles, and the following two requests are given.

\[(10) \quad \text{a. Pon } una \text{ bolita en la caja. b. Pon } unas \text{ bolitas en la caja.} \]

\[
\begin{align*}
\text{Put a/one. SG marble. SG in the box.} \\
\text{‘Put a/one marble in the box’}
\end{align*}
\]

\[
\begin{align*}
\text{Put some. PL marbles. PL in the box.} \\
\text{‘Put some marbles in the box’}
\end{align*}
\]

Regardless of what kind of theory of the plural/singular morphology we choose, the following two requests receive different interpretations in the adult language in a setting where there is a box with marbles and the adult has to act-out in response to one of the following requests.

In a theory in which \textit{una bolita} ‘a marble’ has the feature for number [+sg], defined as in (6) and \textit{unas bolitas} is marked [−sg], the DP marked [+sg] should be interpreted as referring to an atom, and \textit{unas bolitas} should be interpreted as referring to a plurality. No matter which theory we pick, we expect that if children have mastered the distinction, they should do perfectly in an act-out task of this type and provide an atom for (10a) and a plurality for (10b). Even if we assume a privative system for the semantics of plurality in which only singular has a well-defined semantics and plural is interpreted as non-atomic by implicature (Sauerland, 2005, for example), we expect that all younger children (Mexican and Chilean) will have problems of the same sort when compared to older children and adults. Under the Grammar Competition Hypothesis we expect Chilean children to have two grammars with different weights or to have already decided on a grammar with or without plural morphology. For ChMC children a variable behavior is not unlikely, since both adult and child production are around 50%. For ChWC children we expect that the grammar without the plural/singular distinction will carry more weight.

Because the variant [h] is the most frequent among speakers, it is important to verify for Chilean children whether their behaviors when tested in [h] or [s] show any differences. One could argue that children do not have [s] as a plural marker but rather [h], given its prominence in the input (Miller, 2007).

4.1. Subjects

91 children were invited to participate but only 89 children participated in the two experiments as 2 did not pass the warm-up task. In Experiment 1 the plural morpheme was always pronounced as [s]. In this experiment 19 MexWC (4.7–6.5), 36 ChWC (4.7–7.3), 34 ChMC (4.8–8.2) children and 10 Chilean adults (5 ChWC and 5 ChMC) and 6 Mexican adults (3 MexWC and 3 MexMC) participated. Experiment 2 was identical to Experiment 1 except that the native Spanish-speaking researcher always pronounced the plural morpheme as [h]. Experiment 2 involved 11 ChWC
and 9 ChMC (5.1–6.1) children. The Chilean children were recruited from preschools and daycares in Punta Arenas, Chile and the Mexican children from a daycare in Mexico City. All ChWC, ChMC and MexWC adult subjects were the parents of children who participated in this study. Three additional MexMC adult subjects were interviewed to ensure that MexWC and MexMC adult comprehension of the plural morpheme did not differ. In order to examine developmental patterns, children were divided into two age groups in the version of the experiment where the researcher pronounced the plural morpheme as [s], an older and younger age group. It is important to note that none of the younger groups had any formal reading instruction. While we were able to match the younger [s] groups by age, the older groups are not matched in age exactly. The Mexican group is 4–13 months younger than the older Chilean groups and, unlike the Chilean older groups, had not received any formal reading instruction. Table 3 describes the subject pool for Experiments 1 and 2.

### Table 3

Subject pool for Experiments 1 and 2.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Group</th>
<th>Age</th>
<th>Subjects N</th>
<th>Age range</th>
<th>Mean age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 [s]</td>
<td>MexWC</td>
<td>Younger</td>
<td>14</td>
<td>4.7–5.10</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Older</td>
<td>5</td>
<td>6.0–6.5</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td>ChMC</td>
<td>Younger</td>
<td>15</td>
<td>4.8–5.11</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Older</td>
<td>10</td>
<td>6.0–7.3</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td>ChWC</td>
<td>Younger</td>
<td>15</td>
<td>4.7–5.11</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Older</td>
<td>10</td>
<td>6.7–8.2</td>
<td>7.4</td>
</tr>
<tr>
<td>2 [h]</td>
<td>ChMC</td>
<td>–</td>
<td>9</td>
<td>5.1–6.1</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td>ChWC</td>
<td>–</td>
<td>11</td>
<td>4.9–6.4</td>
<td>5.6</td>
</tr>
</tbody>
</table>

(4.9–6.4) and 9 ChMC (5.1–6.1) children. The Chilean children were recruited from preschools and daycares in Punta Arenas, Chile and the Mexican children from a daycare in Mexico City. All ChWC, ChMC and MexWC adult subjects were the parents of children who participated in this study. Three additional MexMC adult subjects were interviewed to ensure that MexWC and MexMC adult comprehension of the plural morpheme did not differ. In order to examine developmental patterns, children were divided into two age groups in the version of the experiment where the researcher pronounced the plural morpheme as [s], an older and younger age group. It is important to note that none of the younger groups had any formal reading instruction. While we were able to match the younger [s] groups by age, the older groups are not matched in age exactly. The Mexican group is 4–13 months younger than the older Chilean groups and, unlike the Chilean older groups, had not received any formal reading instruction. Table 3 describes the subject pool for Experiments 1 and 2.

### 4.2. Method

The task procedure for both Experiments 1 and 2 is very simple and very closely matches children’s everyday language experiences. In fact, Ferenz and Prasada (2002) showed that English-speaking children as young as 1.9 years of age had very little difficulty with this task. Subjects were presented with sets of miniature toys (6 toys per set) followed by 4 singular and 4 plural targets with the indefinite singular (2 feminine and 2 masculine) and asked to give a marble or some marbles, as illustrated in (10). The noun phrases under investigation were always in object position, rather than in subject position (where they would trigger subject–verb agreement). Chilean children were tested either with the plural morpheme pronounced as [s] (Experiment 1) or as [h] (Experiment 2), while Mexican children were only tested on [s] (Experiment 1). Controls involved the quantifiers muchos (many.PL), pocos (few.PL), todos (all.PL), the numbers dos (two) and tres (three). We also checked their ability to count to seven.

The experimenter started the testing session by placing 7 miniature puppies in a row in front of the child and asking the child to count the puppies. Children were then asked to place two puppies (dos perritos) and then three puppies (tres perritos) in a small box. Only two children did not pass this warm-up phase. Next, the experimenter presented each child with two sets of objects (autos ‘cars’ and bolitas ‘marbles’) and told each child: “Te voy a decir cuantas bolitas y cuantos autos tienes que poner en la caja y tu pones la cantidad que yo te digo, te parece?” (**“I’m going to tell you how many marbles and how many cars you have to put in the box and you put in the quantity that I tell you, sound good?”**). The experimenter then began running the main section of the experiment, which involved target and control conditions. The test was administered the same way for adults; however, they were not tested on their ability to count, nor on number terms or controls.

---

6 The comparison between feminine and masculine is important because while the difference between feminine singular and feminine plural is just the /–s/, for the masculine the difference between singular and plural includes a vowel, as illustrated below:

(i) a. un vs unos
   - INDEFINITE-MASC.SG VS. INDEFINITE-MASC.PL
b. una vs unas
   - INDEFINITE-FEM.SG VS. INDEFINITE-FEM.PL

A within-subjects design was used. There were four trials of each target sentence and four trials of each control sentence, making a total of 32 experimental sentences. All experimental sentences were presented in pseudo-random order and sentences involving the same determiner never immediately followed each other.

4.3. Results

Table 4 shows the percentage of ‘more than one’ responses (placing more than one object in the box) in the control conditions for the YOUNGER CHILDREN and the OLDER CHILDREN (see Table 3).

The results in the control conditions show that all children do well. In the *pocos* condition, a plural set was not always chosen but this is not an incorrect answer in this context (*pocos* allows for a reading of just one), although there was a tendency for children to give a plural set to the experimenter. A one-way ANOVA showed no significant differences between the three younger child groups \(F(2,42) = .147, p = .864\), nor the three older child groups \(F(2,24) = .052, p = .949\), nor between the ChWC and ChMC children tested on [h] \(F(1,19) = 1.084, p = .312\). While children performed similarly in the control conditions, they did not perform in the same way in the target conditions. Table 5 shows the percentage of correct responses in both Experiment 1 and Experiment 2.

The results show Chilean children behaving as if they are ignoring plural morphology a fair amount of time. A one-way ANOVA showed significant differences between the three child groups in the plural condition when the plural morpheme was pronounced as [s] \(F(2,41) = 4.902, p < .05\) and also when the plural morpheme was pronounced as [s] for Mexican and [h] for Chilean children \(F(2,31) = 7.625, p < .01\). Post hoc Bonferroni tests showed that the only younger-[s] groups that behaved significantly different from adults were the younger-[s] ChWC children \(p < .001\) and younger-[s] ChMC children \(p < .05\). Mexican children did not behave significantly different from adults \(p = .434\). Turning now to the older groups tested on the variant [s] (older-[s]), a one-way ANOVA showed no significant differences between the older-[s] ChWC, older-[s] ChMC, older-[s] MexWC and Adult groups \(F(3,40) = 1.890, p = .148\). To determine whether there was an age effect, the younger-[s] children were compared to the older-[s] children. The results show an age effect for the ChWC children \(t(1,23) = -2.418, p < .05\) but no age effect for ChMC children \(t(1,23) = -1.629, p = .117\) and no age effect for MexWC children \(t(1,17) = -1.137, p = .271\).

A comparison between feminine and masculine items showed no effect for the ChWC younger group \(t(1,13) = -1.472, p = .165\) nor for the younger ChMC group \(t(1,13) = .366, p = .720\), nor for any of the other groups.
Finally, if we look at the ChWC and ChMC tested on [h], we find no significant differences between these two groups ($t(1,18) = -0.433, p = .633$). Child responses were not compared to adult responses given that we did not test adults on [h].

In order to determine whether the same children always treated the plural indefinite *unos* as plural or whether behavior was variable for each child, we examined the systematicity in response patterns on *unos* for each child. We are defining systematic behavior as having the same response (plural or singular) in at least 3 out of 4 trials. Variable comprehension would be having a singular response for 2 of the 4 trials and a plural response for the other 2 of 4 trials. The results are given in Table 6.

The results are striking. Most young ChWC children did not associate the plural morpheme to ‘more than one’, and only about half of the ChMC children did. Interestingly, for the older groups, we find significant improvement. The Production hypothesis can be quickly eliminated. Chilean children are not performing as if they all have a grammar with plural morphology.

The differences between the younger Mexican Children on one hand and the younger Chilean Children are very clear and support the overall hypothesis of the Variational model, namely that the more variability/ambiguity in the input the longer it will take the learner to converge on to the adult grammar. In other words, while Mexican children by 5 years of age are adult-like, ChMC and ChWC take longer to reach an adult-state. Furthermore their behavior both as a group and individually correlates with the rate of omission in the adult free speech data, as summarized in Table 7 below. Interestingly, however, we do not find many children behaving as if they have two grammars competing, since this would lead us to expect a much more variable behavior in comprehension for each individual child, which we did

### Table 6
Percentage of systematic responders for *unosunas*.

<table>
<thead>
<tr>
<th>Group</th>
<th>% of systematic responders (3–4 plural responses on <em>unos</em>)</th>
<th>% of systematic responders (0–1 plural responses on <em>unos</em>)</th>
<th>% of total systematic responders</th>
<th>% of variable responders (2 plural and 2 singular responses on <em>unos</em>)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plural pronounced as [s]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Younger</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MexWC</td>
<td>86</td>
<td>14</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>ChMC</td>
<td>53</td>
<td>33</td>
<td>86</td>
<td>13</td>
</tr>
<tr>
<td>ChWC</td>
<td>27</td>
<td>67</td>
<td>94</td>
<td>6</td>
</tr>
<tr>
<td>Older</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MexWC</td>
<td>100</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>ChMC</td>
<td>90</td>
<td>10</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>ChWC</td>
<td>70</td>
<td>20</td>
<td>80</td>
<td>10</td>
</tr>
<tr>
<td>Plural pronounced as [h]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChMC</td>
<td>22</td>
<td>56</td>
<td>78</td>
<td>22</td>
</tr>
<tr>
<td>ChWC</td>
<td>27</td>
<td>73</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>
not find. Rather we find most ChWC children ignoring the plural morphology, which is compatible with the hypothesis that their grammar has no plural morphology.

5. Discussion

We found that, while some Chilean children systematically interpreted plural morphology as an indication of ‘more than one’, most systematically ignored plural morphology. Overall, it seems that mastery of lenition is not a reasonable explanation for the combined production and comprehension results. An alternative explanation for the results based on the idea that the plural interpretation is dependent on children’s ability to calculate implicatures is also untenable. If that were the case, we should expect development in all groups and we should not find dramatic differences between groups.

The results support the Grammar Competition Hypothesis. Production results clearly showed that children who were exposed to more overt realization of plural morphology produced more plurals. Mexican children always produced the plural and their comprehension was not significantly different than the adult comprehension. ChMC children produced much more plural morphology than ChWC children. In comprehension, we again found that Chilean children who were exposed to more reliable input with respect to plural tended to prefer a grammar with the distinction. ChWC children, on the other hand, preferred the grammar without plural/singular distinction and consequently ignored the overt marking of the plural morpheme in comprehension tasks. What we also found was that the [h] form, in spite of being more frequent, did not help most children. This is not particularly surprising if we take into consideration the fact that [h] is acoustically less salient. According to Fox (2006), [s] has a much longer duration median than [h]. The median value for [s] is 80 ms, while the median value for [h] is 50 ms.

What we did not find was children using both grammars in the two experiments and exhibiting a variable response, which we could have predicted given that at least for ChWC children the rate of omission for both children and adults was close to 50%. The lack of variability within individuals may be related to the fact that children are using the most rewarded grammar within the experiment. But the use of the grammar that does not make an overt distinction between singular and plural does not per se explain why children systematically gave singletons to instructions with unas but no plurals with una.

There is however a possible explanation for the bias towards singular reading of the indefinite una and unas, namely the fact that the indefinite article un/una is homophonous with the number for ‘one’ in Spanish. Children who know that the indefinite is associated with ‘one’ yet do not associate the plural marker to ‘more than one’ may fulfill the experimental instruction by with a singleton.

It remains open for exploration the question of exactly what makes plural morphology so unreliable for ChWC children. Is it sheer overall omission of number marking or is it that number is missing in some crucial places? What is clear, though, is that variation that causes ambiguity does make children take longer to acquire the ability to use variable forms in comprehension.

References


Cantú-Sánchez, M., Grinstead, J., 2004. Nominal number and gender agreement in child Spanish SLI. Poster presented at Generative Approaches to Language Acquisition in North America (GALANA), December 19, University of Hawai’i at Manoa.


Ferenz, K., Prasad, S., 2002. Singular or plural? Children’s knowledge of the factors that determine the appropriate form of count nouns. Journal of Child Language 29, 49–70.


