

The Acquisition of Spanish Gender by
English-speaking Children in a
Partial Immersion Setting

Donna Bosworth Andrews

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

2004

Program Authorized to Offer Degree:
Department of Linguistics

UMI Number: 3139443

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 3139443

Copyright 2004 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346


University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Donna Bosworth Andrews

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Co-Chairs of Supervisory Committee:

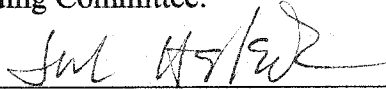


Julia Herschensohn



Carol Stoel-Gammon

Reading Committee:



Julia Herschensohn



Carol Stoel-Gammon



Jurgen Klausenburger

Date 6/11/04

In presenting this dissertation in partial fulfillment of the requirements for the doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to Proquest Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, to whom the author has granted the "right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature Donald Anderson

Date 6/11/04

University of Washington

Abstract

The Acquisition of Spanish Gender by
English-speaking Children in a
Partial Immersion Setting

Donna Bosworth Andrews

Chairs of the Supervisory Committee:
Professor Julia Herschensohn
Department of Linguistics
Professor Carol Stoel-Gammon
Department of Speech and Hearing Sciences

This study investigates how English-speaking 5-7 year olds in a Spanish immersion program acquire the Spanish gender system.

Perez-Pereira (1991) has shown that native Spanish-speaking children do not attend to semantic cues (reference to a male or female entity) nearly to the extent that they attend to phonological cues (word endings). Similar research by Karmiloff-Smith (1979), Levy (1983), and Mills (1986) support Perez-Pereira's findings that children attend to language-internal cues to grammatical gender, not language-external cues.

However, learners of second languages may use different strategies in acquiring gender systems than first language learners do. Taylor-Brown (1984) showed that English-speaking 3rd-9th graders in a Canadian French immersion program were unable to use French phonological cues in the same way that native French-speaking children do. Instead, they relied on consciously employed "rules of thumb" to remember noun genders. The question here is whether younger

children follow the acquisition pattern of their older L2 counterparts, or of L1 learners of +gender language

In this study, kindergartners and first graders in a Spanish immersion program looked at pictures of people and animals, and listened to two puppets using masculine and feminine articles to refer to the entities in the pictures. They were then asked to decide which puppet was using the correct article. Results indicate that kindergartners and boys were more successful at correctly assigning gender when they heard a morpho-phonological cue that corresponded to the correct gender than when they only saw a semantic cue to grammatical gender. Results also show that children were better at assigning gender when they had two converging cues to gender, rather than a single cue. This suggests that, like L1 learners of a +gender language, the kindergartners and boys use the linguistic gender cue to a greater extent than they use the semantic (sex) cue, and to a greater extent than first graders and girls. These results contribute to our understanding of how age, cognitive development, and first language may influence acquisition of a +gender L2.

Table of Contents

	Page
List of Figures	iii
List of Tables	iv
Chapter 1	1
1.1 Introduction	1
1.2 The Critical Period Hypothesis	4
1.3 The role of the native language in L2 acquisition	13
1.4 Conclusion	17
Chapter 2	18
2.1 What is gender?	18
2.2 What's gender good for anyway?	22
2.3 The gender system of Spanish	25
2.4 The gender system of English	28
2.5 Psychological correlates of gender	30
2.6 Conclusion	34
Chapter 3	36
3.1 Acquisition of gender in L1	36
3.2 They "syntax first" approach	39
3.3 Gender acquisition research supporting "syntax first"	46
3.3.1 French	47
3.3.2 Hebrew	57
3.3.3 German	61
3.3.4 Spanish	65
3.3.5 Icelandic	69
3.4 Second language acquisition of gender	69
3.5 Current research questions	78
3.6 Conclusion	79
Chapter 4	81
4.1 Subjects	81
4.2 Design	84
4.3 Materials	91
4.4 Procedure	92
4.5 Analysis	96
4.6 Conclusion	99
Chapter 5	100
5.1 Introduction	100
5.2 Use of linguistic vs. natural cues to gender	101
5.3 Statistical analysis of complete data	103
5.4 Do kindergartners and first graders make different use of cues?	107
5.5 Do boys and girls make different use of cues?	111
5.6 Do children overgeneralize one gender?	114
5.7 Conclusion	117
Chapter 6	118

6.1 Single cues vs. multiple cues	118
6.2 Age- and sex-related differences	119
6.3 Familiarity with test stimulus.....	122
6.4 Generic use of masculine forms in Spanish.....	123
6.5 Avenues for future research.....	124
Bibliography	128
Appendix A: Pictures used in the experiment.....	135
Appendix B: Proportion right per child, by condition	137
Appendix C: Statistical tests	141

List of Figures

Figure Number	Page
3.1: Network of verb correspondences	41
3.2: Network of adjective correspondences	42

List of Tables

Table Number	Page
3.1: Cues tested in Karmiloff-Smith	49
4.1: Combinations of cues to grammatical gender.....	85
4.2: Words and gender cues in Phase I	88
4.3: Words and gender cues in Phase II.....	89
4.4: Example of calculations by condition.....	98
5.1: Means correct for all groups and conditions.....	102
5.2: Differential use of cues (all children)	104
5.3: Differential use of cues statistics (all children)	105
5.4: Kindergartners and First Graders mean correct by condition.....	107
5.5: Kindergartners differential use of cues.	108
5.6: First Graders differential use of cues.....	110
5.7: Boys' and girls' mean correct per condition.....	112
5.8: Boys' differential use of cues	112
5.9: Girls' differential use of cues.....	113
5.10: Means correct for masculine nouns vs. feminine nouns.....	115
5.11: Statistical difference for masc. and fem. nouns, by condition.....	116
6.1: Means correct for all groups and conditions.....	127

Acknowledgements

Although the writing of this dissertation was a solitary endeavor, it is nonetheless a collaborative effort made possible by the support and generosity of many, a few of whom I must acknowledge.

My advisors, Julia Herschensohn and Carol Stoel-Gammon, were ever positive and encouraging, keeping me on course despite my numerous attempts to become derailed. Without their clear guidance and outright refusal to let my determination flag, this dissertation might never have been completed.

I would also like to acknowledge Philip Dale for introducing me to the field of child language acquisition early in my graduate career. His interest in my first attempts at child language research helped me achieve the confidence to get this far. Also thanks are due to Paul Aoki of the UW Language Learning Center, for encouraging me to look for ways to include the newly-opened John Stanford International School in my research plans. His direction and involvement helped shape my earliest ideas for this dissertation and brought me into contact with those whose permission and support were essential.

In that light, I would also like to thank Karen Kodama, Principal of the John Stanford International School; Maria Buceta-Miller and Dolly Morales, kindergarten and first grade teachers at JSIS; and the Seattle School District for sanctioning and supporting this research. And not least, to the children who participated in this research and the parents who allowed them to do so, I also extend my gratitude.

A special word of thanks also goes to Lola Rodriguez, who, during her visiting scholar year at the UW, lent her support, her native Spanish, and her positive and cheerful demeanor to every aspect of this study. Her assistance with conducting the Spanish portions of this experiment made the study possible. I must also acknowledge the help I received from the UW Statistics department, especially Bryan Shepherd and Jay Jones.

Finally, and most importantly, the biggest thank you of all goes to my husband Rich, my parents, and the rest of my family, who allowed me the time and space to write this dissertation, often to their own personal sacrifice. To Ben and Declan, my amazing children and first language acquisition subjects, who have done without me so often, I promise never do to this again.

Chapter 1

1.1 Introduction

Research in second language acquisition (L2A) is multifaceted, but overall, the fundamental question which characterizes all L2A research is this: to what extent, and in what ways, is the acquisition of a second language the same as or different than learning a first language? We might expect to find differences for one very obvious reason: the initial state of the learner, in terms of age, cognitive ability, and linguistic ability, is different in L1A than in L2A. The L1 learner is a baby where the L2 learner may be anyone from a very young child¹ to an adult. Because the L1 learner is a baby, his or her cognitive abilities are those of a baby. There is no way for babies to strategize or plan how to learn their L1, as their level of cognitive competence does not allow it. An L2 learner could, however, bring his or her considerable cognitive abilities to bear on their L2, possibly even approaching it as they would approach any other learning, such as science or mathematics. Finally, L1 learners by definition have no linguistic competence outside of the L1 itself; their total linguistic competence can be assessed within the L1. The L2 learner, by contrast, already has an L1 and therefore has a level of linguistic competence that is independent from the L2 being learned. A very real issue within L2A research is whether and how these

¹ Not included in the definition of L2 learners are babies who learn more than one language simultaneously from birth. In such cases it is debatable whether either language could be classified as the "first" or "second" language. The complexities of the issues surrounding bilingualism are largely irrelevant to the research in this dissertation, and therefore nothing will be said about it here.

differences between L1 and L2 learners affect and shape the ways in which L2 learners approach L2A.

In this dissertation, the topic through which we shall examine such issues as those above is the topic of grammatical gender. It has been widely shown that L1 learners of languages with grammatical gender ignore potential semantic clues to grammatical gender of nouns (Karmiloff-Smith, 1979; Levy, 1983a & b; Perez-Pereira, 1991). That is, despite the fact that grammatical gender often has natural correlates with sex in the real world (e.g., females and males fall into different gender classifications) these natural correlates have little bearing on a child's acquisition of grammatical gender. The importance of gender research in first language acquisition (L1A) is in showing that children's earliest linguistic classifications are based on structure rather than on a form-meaning relationship.

The same has not been as conclusively shown for L2 acquisition. In addition, most L2 research focuses primarily on adult learners (Carroll, 1989; Franceschina, 2003; Hawkins 1998; White et al. 2004) which begs the question of whether adult and child L2 learners should be studied together or distinctly. There are good reasons to suspect that child learners of a second language may use different language learning strategies than adults do. One reason is the oft-cited observation that people who begin learning an L2 in childhood generally attain a higher level of proficiency than those who begin later in life (Johnson and Newport, 1989), especially in the area of accent (Patkowski, 1990). Johnson and Newport offer their research as evidence for a "critical period" for second language learning. However, it has also been suggested that the differences in

eventual L2 attainment between children and adults can be attributed to different learning strategies or cognitive abilities at different ages, rather than a true “critical period” (Flege, 1987). Adults may bring different learning strategies to bear on language learning because of their years of experience in learning environments and in “being taught”, whereas children are less likely to overtly strategize methods for learning, but rather just absorb the input available to them. For example, with respect to gender, it may be that adult learners use strategies such as categorization (female things take feminine gender, male things take masculine gender) that are too cognitively advanced for children, forcing children to rely on linguistic factors that adults overlook.

Whatever is the case, it is certainly true that there has been very little research on child L2 learners acquiring a +gen L2 but whose L1 is -gen. An important question, then, is whether children who have already learned L1 to a high communicative level (5- and 6-year-olds), and therefore have formed many form-meaning correspondences, continue to employ a purely linguistic strategy in their language learning or opt instead for a strategy which takes meaning from the world they already know and looks for the linguistic forms to express it.

The present dissertation is an attempt to contribute to the question of how grammatical gender is acquired by looking at English L1 children learning Spanish as L2. By sampling the children’s understanding of gender early in their exposure to Spanish, the study described here investigates whether child L2 learners use the same strategy in acquiring grammatical gender as their L1 counterparts. That is, do kindergartners and first graders with English L1 use

linguistic cues (e.g., phonology and morphology of words) to determine grammatical gender, or do they instead use natural cues (related to male and female sex characteristics) to do so?

This dissertation is organized as follows. The remainder of Chapter 1 will discuss the relationship between L1 and L2 acquisition, focusing specifically on two areas: the Critical Period Hypothesis, and role of the native language in L2 acquisition. Chapter 1 further explains how the experiment detailed in this dissertation sheds light on both of these issues. Chapter 2 introduces the concept of grammatical gender, its purpose and manifestations, and briefly describes the gender systems of Spanish and English. It also briefly lays out the potential psychological relationship between natural and grammatical gender. Chapter 3 discusses the acquisition of gender in L1, placing it within a formalist hypothesis of L1 acquisition, giving us a basis of comparison for the L2 children in this study. Chapter 3 also reviews research of gender acquisition in L2 in order to compare with L1. In Chapter 4, the present study is described. Chapter 5 presents the results of the study. Chapter 6 discusses implications of the results and concludes by relating the findings from this study to the theories discussed in Chapter 1, Chapter 2, and Chapter 3.

1.2 The Critical Period Hypothesis

An issue in L2A is the relationship between age and L2 ability. Much research centers around the critical period (CP) which asserts that there is a

window for acquiring any language to native (in the case of L1) or native-like (in the case of L2) proficiency. The CP was originally proposed for L1 acquisition by Lenneberg (1967) to explain why children who experienced brain trauma to their left hemispheres recovered their language abilities quickly and with apparent ease, while adults with similar brain injuries often did not recover their language at all. Lenneberg suggested that because lateralization of the cerebral functions was incomplete in children until around puberty, language could relocate to an undamaged area of the brain until then. The CP gained relatively wide acceptance with the discovery of Genie, a 13-year-old girl kept in abusive conditions who had never learned an L1 and who, despite careful instruction, never really succeeded in doing so (Curtiss, 1977). While subsequent research has questioned the exact nature of the age constraint (e.g., Krashen (1973) showed evidence that lateralization is largely complete well before puberty) there does seem to be ample evidence that very young children exposed to language will naturally acquire it in ways that older children and adults exposed to language will not.

Studies of deaf children who grow up in the absence of sign language (and therefore, in the absence of a language they can acquire naturally) have also supported the idea of a CP for L1. For example, Newport (1988, 1990) has found that children who grow up in homes where signing is used as a normal means of communication uniformly acquire a full and rich sign language, whereas children who are not exposed to signing until later often develop inconsistent and less native-like competency. Advances in the measurement of neural functions have also contributed to the debate. Functional Magnetic Resonance Imaging (fMRI)

studies of ASL speakers (Newman et al., 2002) showed that there are clear differences in the brains of early and late sign learners. Certain regions of the right hemisphere are active when children who learned ASL before puberty read sign language, but these same regions are not active when children who learned ASL after puberty are engaged in exactly the same activity.

Whatever the evidence for a critical period in L1 acquisition, the evidence for a CP in L2A is more varied and less definitive. One noteworthy study which claimed to show evidence for a Critical Period is Johnson and Newport (1989). Their study of Chinese and Korean L1 learners of English, whose age of arrival (AoA) in the US ranged from 3 to 39, showed that the earlier the age of arrival in the L2 country, the higher the level of proficiency years later. Johnson and Newport classified their subjects into Early Arrivals (arrived in the US before age 15) and Late Arrivals (arrived after age 17) and asked them to give grammaticality judgments on a set of 276 English sentences, half of which were ungrammatical. They found that for the Early Arrivals, AoA was highly correlated to eventual achievement in L2. Native-like attainment was common for L2 learners first exposed to L2 before age 7, but a decline in eventual ability began at about AoA 8, and continued declining in a linear fashion up to AoA 15. There was then an “elbow” effect; Late Arrivals’ eventual level of attainment continued to decline with a less predictable pattern. Johnson and Newport interpreted this to mean that the critical window for L2 acquisition closes around the age of 7, and after that time it is extremely difficult to learn a foreign language to native-like proficiency.

It has been pointed out, however, that if a critical period indeed exists between the ages of 0-7, we should not see proficiency continue to decline as a function of age for up to 10 more years (Bialystok and Hakuta, 1999; Birdsong and Molis, 2001). If a “language learning window” closes at age 7, what is the nature of the changes taking place between the ages of 7 and 16 that allow learning to continue but with less success as a function of age? Such a decline casts doubt on the idea of this period as “critical.” More recent research confirms a decline in ability to achieve native-like competence in L2 correlated with AoA, but the “elbow-effect” that Johnson and Newport found in their data is not shown. L2 attainment appears to decline with age of first exposure well into adulthood, suggesting that a critical period (whether ending at 7, 12, or 16, all of which have been proposed as cut-off ages for learning L2) may not be the reason (Bialystok, 1997; Birdsong and Molis, 2001; Flege, 1999; Kellerman, 1995).

The idea that there is a CP for language acquisition as a whole may suffer from a level of overstatement. Some research has indicated that evidence for a CP depends on which aspects of language are being investigated; for example, syntax, morphology, or phonology. Phonology is the area in which it has been most convincingly shown that earlier exposure is undoubtedly better. Patkowski (1990) re-examined his own (1980) data from 67 subjects whose age of first exposure to English L2 ranged from 5 to 50, which showed that those with earlier AoA had the lowest degree of foreign accent (as perceived by native English speaking judges). In his reanalysis he separated the adult-arrivals from the child-arrivals and showed that there was a great discontinuity between the two groups

in terms of accent, suggesting a major difference between the two groups abilities rather than a gradual decline from one into the other.

Age may play another role in L2 acquisition, affecting the rate of learning. Krashen (1982) reviewed numerous studies and concluded that based on all the evidence, older learners had an advantage early in learning, showing signs of acquiring some structures more quickly than younger children, but that in terms of eventual attainment, learners who were exposed to their L2 at a younger age were more likely ultimately to reach a higher level of proficiency. Snow and Hoefnagel-Hohle (1978) conducted one of the earliest studies to show a relationship between age and rate of learning. Over a period of 10 months, they studied the acquisition of Dutch by L1 English speakers in three age groups: 8-10 year olds, 12-15 year olds, and adults, measuring proficiency at three, six and ten months of naturalistic exposure to the target language. They found that initially, the adolescent group did the best, followed by adults and the younger children (except in pronunciation, where there were no significant differences), but by the end of the study the 8-10 year olds had narrowed the gap considerably.

The early advantage shown by older children may hold only for certain linguistic features, rather than for all aspects of language. Harley (1984) found that after 1000 hours of French immersion, older learners (starting at 13) showed a clear advantage over younger learners (starting at 5 and 6 years old) only in number agreement and SOV word order affecting pronouns. In all other areas of syntax tested, the older and younger students showed no significant differences. Slavoff and Johnson (1995) point out that the early advantage enjoyed by the

older learners in the Snow and Hoefnagel-Hohle study might have been due to the similarities between Dutch and English, which the older learners may have had more ability to rely on (more on the relationship between L1 and L2 below). Slavoff and Johnson therefore studied speakers of Asian languages as L1 (Korean, Vietnamese, Chinese, and Japanese, to be maximally different from English L2) who were between the ages of 7-9 and 10-12 at AoA in the US, and who had been exposed to English through immersion in public school anywhere from 6 months to 3 years. Using the same grammaticality judgment test of Johnson and Newport (1989) mentioned above (shortened to accommodate the attention span of their younger subjects), Slavoff and Johnson looked at the level of proficiency on twelve morphological elements or syntactic structures at 6 months of exposure and again at 3 years of exposure. They found that rate of acquisition was not predicted by AoA, but rather by length of immersion in the target language. That is, the longer the learners were exposed to the L2, the more quickly they acquired new L2 structures.

Some researchers have proposed theories of why “younger is better” that do not involve a CP, but rather take advantage of other correlates of age. Newport (1988, 1990) suggests that young children’s attention span and cognitive ability to attend to only small amounts of input at a time makes them ideally suited for morphological analysis, since morphemes are chunks of linguistic data that correspond to the amount of information small children are capable of taking in. She reviews her own research on ASL learners, in which she examined the learning patterns of “native/native signers” (those who learned ASL as their first

language from parents who themselves are “native signers”), “native/late signers” (those who learned ASL as their first language from parents who were not themselves native signers) “early learners” (those who learned sign beginning age 4-6 from peers at a residential deaf school), and “late learners (those who learned sign beginning age 12 from peers at a residential deaf school). The pattern which emerges in comparing sign acquisition across these groups is that the two native groups show strong evidence of morphological analysis, gradually incorporating more and more morphemes into a sign, at first sequentially and later simultaneously. Late learners, on the other hand, show evidence of learning signs holistically rather than as morphological units. Early learners showed more evidence of morphological analysis, but not as much as native learners. Most interesting, native/native and native/late signers were virtually indistinguishable in their use of morphological analysis, suggesting that even when the parents’ input contained many holistic forms (that is, forms which could not themselves be morphologically analyzed) their children imposed a morphological analysis on them.

Newport points out that, amid very young children’s only limited capabilities in any other cognitive area (such as memory, vision, or ability to generalize, for example), it is inconsistent to posit that in language alone, very young children possess superior skills to older humans. It makes more sense to posit that “the very limitations of the child provide the bases on which to perform successful language learning, and that indeed without these limitations... the acquisition of some aspects of language is more likely to fail” (Newport 1980,

165). One way this may happen is that children may at first store the smaller morphemes as units of meaning rather than anything morphologically complex, facilitating an increased awareness of these morphemes when they are encountered in complex contexts (i.e., words). While Newport is quick to assert that her data on ASL acquisition do not require such a hypothesis, it is at least as explanatory, and more consistent with children's other abilities, than posting a dissociation between children's linguistic abilities and their limited abilities in everything else. Cochran, McDonald and Parault (1997) also note that as language is broken down into smaller units, there will be fewer units to map to meaning, reducing the complexity of figuring out the form-meaning relationship (see Ch. 2). Once this mapping is figured out, children can use and recombine the forms to create grammars.

A different theory (cited in Flege, 1999) of L2 acquisition is one he terms the "exercise hypothesis" in which "one's ability to learn to produce and perceive speech remains intact across the life span, but only if one continues to learn speech uninterrupted." He points out, however, that this would be nearly impossible to test, given the paucity of participants who continue language learning endeavors uninterrupted throughout their lifetime.

Research on L1 development has shown that children acquire various grammatical features of their L1 at different developmental stages; for example in English, passives are acquired much later than, say, question forms. In +gen languages, gender is an early-acquired structure, generally completely acquired by the age of 3 (Levy, 1983a & b; Mills, 1986; Perez-Pereira, 1991). It is possible

that child L2 learners of a +gen language may have an advantage over adults simply because children are still paying attention to linguistic cues as they acquire complex structures in their own language, whereas adults have long since stopped the process of acquiring and therefore are less attuned to the linguistic cues.

Werker and Tees (1984) and others have showed that babies can distinguish all the potential sound differences between phonemes up until about 8 months of age. At about 10 months, babies lose the ability to distinguish contrasts which are not phonemic contrasts of their language environment. The theory is that because the rest of the distinctions are irrelevant babies begin to concentrate their effort on those sounds that matter in their language and ignore those that don't. However, the ability to hear those "irrelevant" distinctions remains latent and can be evoked by minimal amounts of natural exposure to another language in which the distinctions are manifested, and children can continue to hear those distinctions as long as their ability to do so continues to be exercised.

Because gender is such an early-acquired feature in a +gen L1, it is a good candidate to test with young L2 learners. Following Newport's (1988, 1990) hypothesis, it may be that gender is acquired better by younger learners because it is encoded in very small morphemes (one or two phonemes) which are rarely more than a word or two away from each other. Young learners may be uniquely suited to attend to exactly this type of linguistic input in a way that older learners attending to greater amounts of input are not.

Age may interact with another variable, that is, the way in which learners use their L1 as a springboard for their L2. It is to the role of the native language in L2 that we now turn.

1.3 The role of the native language in L2 acquisition

As mentioned above, L2 learners bring their L1 with them to the task of learning L2. This is not debatable; it is neither expected nor possible that learners can ignore or “shut off” L1 while learning L2, since they need to function linguistically and have (at least initially) only their L1 through which to do that. However, the extent to which and manner in which L2 learners use their L1 in their L2A is a matter of considerable debate. Early theories of L1 influence in L2 learning were rooted in the Contrastive Analysis Hypothesis (Lado, 1957) which stated that “those elements that are similar to (the learner’s) native language will be simple for him, and those elements that are different will be difficult” (Lado, 1957; 2). Since the state-of-the-art theories of language learning were at the time based in behaviorism, it was logical to assert that learners “transferred” their linguistic behavior from L1 to L2. Where the two languages were similar, transfer would facilitate acquisition, but where the languages were different, transfer (also referred to in this case as “interference”) would hinder language acquisition (Lado, 1957; Weinreich, 1953).

It is a legendary story in the field of L2A how Contrastive Analysis failed to predict where learners would fail and succeed at language learning. Its

association with behaviorism naturally hastened its downfall, and for many years any discussion at all of “transfer” in acquisition studies was dismissed as unscholarly and, by definition, wrong. The past 20 years though have seen a renewed acceptance of the influence that the L1 has on L2, usually under the more neutral term “cross-linguistic influence” (Sharwood-Smith and Kellerman, 1986). Larsen-Freeman and Long (1991) point out that cross-linguistic influence in current research is not limited to “transfer” or “interference” but also accounts for phenomena such as “avoidance, borrowing, over-production and L2-related aspects of language loss” (Larsen-Freeman and Long, 1991; 97).

Instances of transfer in the sense the term was first used (making use of similarities in the L1 to facilitate acquisition in L2) are of course difficult to establish. If a feature that is present in both languages is acquired early, there is nothing to say whether that structure was acquired due to transfer, or due to some language acquisition process at work in the act of learning, or possibly due to a combination: transfer facilitating some linguistic process. Before the present study began (after about 2 months exposure to Spanish), for example, the children were observed in their classrooms to use code-mixing in addressing their Spanish-speaking teachers which included dropping fully formed Spanish determiner phrases (DPs) into English sentences (e.g., “write it on *la pizarra*” (the chalkboard) or “I have to go to *el baño*” (the bathroom)). This shows that the children had already acquired a DP category for Spanish. What is unclear is whether the similarity of English DP to Spanish DP, especially as pertains to requiring articles in the singular, facilitated this acquisition, or whether this early

acquisition was the result of a natural developmental pattern in language acquisition.

Gender provides another avenue to explore the influence of L1, however in this case English and Spanish are not similar. The children's L1 English contains no noun-class category equivalent to gender, and therefore no syntactic agreement based on anything analogous to gender. These children will ultimately need to acquire a feature in their L2 which effectively does not exist in their L1, except in a very limited fashion in 3rd singular pronouns (he/she, him/her, his/her) where the feature is entirely correlated to natural sex. The question is whether the lack of a gender category in the children's L1 will somehow hinder their acquisition of the category in L2. Bruhn de Garavito and White (2000)² investigated this question by comparing their own data on gender agreement accuracy with Hawkins (1998) results. Bruhn de Garavito and White assessed gender agreement accuracy in L1 French speakers learning Spanish (where both L1 and L2 are +gen languages with a masculine/feminine gender distinction), whereas Hawkins assessed the accuracy of English L1 students learning French (where L1 is -gen but L2 is +gen). The comparison showed that both groups had similar levels of accuracy despite the fact that Bruhn de Garavito and White's subjects had substantially less exposure to the L2 (1-2 years classroom instruction as opposed to Hawkins' subjects who had participated in an immersion program and had as much as 10 years exposure to the language). Thus it is possible that the

² This study will be reviewed in greater detail in Chapter 3.

presence of the gender feature in L1 led to an accelerated rate of acquisition of gender in the L2.

There are three possible ways in which the children in the present study might approach the new gender feature in their L2 (the list below should not be interpreted as a suggestion that the children's eventual attainment will be constrained by the L1 in any of these ways; these are hypotheses regarding the manner in which the feature is acquired at the beginning of their L2 experience, not their ultimate proficiency in the language):

1. They might overlook the feature entirely and therefore not acquire gender at all (L1 influence = avoidance).
2. They might posit a 1:1 correspondence between grammatical gender and natural sex (in effect, transferring the limited correlation of grammatical gender with natural gender from their L1) and acquire gender incorrectly (L1 influence = interference)
3. They might acquire the feature in the same manner as Spanish L1 children, through attention to the morpho-phonological noun endings, thereby acquiring gender correctly (little or no L1 influence)

The first two of these possibilities involve L1 influence on the learners' interlanguage whereas the third possibility is that the lack of an analogous feature in the L1 will not affect the acquisition of the L2 feature. We have already seen that children did not avoid producing the gender features in their DPs, although prior to the study there was no way to determine a priori whether they had

memorized “chunks” of Det+Noun and used these as mentioned above. This study will help illuminate whether the children referred to their limited gender category in English as an aid in acquiring Spanish gender, or whether the lack of a category exerted no influence in the acquisition of this feature.

The role of the L1 in L2 learning has been extensively explored in adult L2 acquisition (e.g., Gass and Ard, 1984; Kellerman, 1978; Schachter, 1974; Zobl 1980), as well as to some extent in child L2 acquisition (e.g., Hernandez-Chavez, 1977; Huang, 1970; Keller-Cohen, 1979) although it is not at all uncommon for researchers to make reference to native language features when they appear to be relevant to the observed acquisition patterns, regardless of whether this is the main focus of the study. This study will address the role of the minimal gender feature in these children’s L1 English in the acquisition of their L2 Spanish.

1.4 Conclusion

This chapter has set the stage for the study to be reported. The nature of the research question has been presented, with special reference to the Critical Period Hypothesis and the potential influence that L1 may have on the acquisition of L2. The next chapter will explore the concept of grammatical gender more fully, both in Spanish and in other languages, as well as the limited English gender feature. Chapter 2 also highlights some of the psychological correlates of gender, which may be relevant for the questions addressed in this study.

Chapter 2

Before presenting the experiment or its theoretical underpinnings, it is important to understand the nature of the linguistic phenomenon being examined. This chapter will explain what a “grammatical gender system” is, and what role it plays in the structure of the language. It will then give a brief overview of the gender systems of Spanish and English, and finally, review some recent experiments that look at the potential interaction between semantic and formal gender.

2.1 What is gender?

According to Hockett (1958, 231, quoted in Corbett, 1991; 1) "genders are classes of nouns reflected in the behavior of associated words." That is, the gender of a noun may dictate the form of the words associated with it, especially determiners, adjectives, quantifiers and verbs (occasionally adpositions and adverbs as well, although gender agreement with these lexical types is much rarer (Corbett, 1991) and does not apply to Spanish, so shall not be discussed here). Thus agreement is the diagnostic for gender. Maratsos and Chalkley (1980), in fact, make the stronger claim, that it is agreement which defines gender classes.

For nouns, gender is a property of the word itself and as such it is part of the noun's entry in the mental lexicon. This is sometimes referred to as “inherent

gender” (Carroll, 1989; 1999; or “controller gender” (Corbett, 1991)). Inherent gender is a feature of the noun which dictates the agreement relationships that the noun can enter into with the other constituents of the DP of which it is the lexical head, namely determiners and adjectives. Inherent gender is to be distinguished from “derived gender” (Carroll, 1989; 1999); also called “target gender” (Corbett, 1991). The classes of words that manifest agreement, (in Spanish, adjectives and determiners) vary in form depending on the gender of the noun they are associated with; that is, their gender is not a part of their lexical entry but is derived from agreement with the head noun through the morpho-syntactic properties of the language (Carroll, 1989; Corbett, 1991).

Under the minimalist program (Chomsky, 1995), gender is a ϕ feature found on the head noun. The purpose of ϕ features is to check agreement with the corresponding features attached to associated words. The inherent gender of the head noun is said to be “interpretable” in the sense that it carries meaning for semantic interpretation. The derived gender of the associated adjectives and determiners is “uninterpretable” in the sense that they do not contribute to the semantics of the sentence, but are there only to check association with the noun. Other ϕ features are number, case, and tense.

While English speakers may generally think of gender as a binary distinction (masculine / feminine) or ternary distinction (masculine / feminine / neuter) this is by no means always or even usually the case across languages. Some languages distinguish genders along the lines of animate / inanimate, others along the lines of human / animal / inanimate, and others have many more word

classes (for instance, some Swahili and other Bantu languages may have as many as 16 genders). The point to be made here is that while "gender" in English corresponds in a very real way to natural sex, grammatical gender should not be identified only with sexual characteristics.³

In fact, it often seems that not only is grammatical gender not identified with sexual characteristics, it is completely arbitrary and unrelated to natural sex. Why, for example, should a "floor" be masculine but a "wall" be feminine, as is the case in Spanish (*el suelo, la pared*)? Nor does there seem anything inherently different in terms of natural gender for a knife, fork and spoon which in German are neuter, feminine and masculine respectively. In addition, it is often pointed out that many languages (the major European languages among them) vary greatly in which gender is assigned to any given noun. For example, "banana" is masculine in Spanish (*el plátano*) but feminine in Greek (*e banána*); "apple" and "orange" are feminine in Spanish (*la manzana, la naranja*) but neuter in Greek (*to mílo, to portokáli*); "house" is masculine in Russian (*dom*), feminine in Spanish (*la casa*) and neuter in Greek (*to spíti*). Thus assignment of non-sexed nouns to particular genders is semantically arbitrary.

³ The term "gender" is often used interchangeably with "declension type" but Corbett (1991) makes the point that the two are not synonymous. While it is true that declension types often correlate strongly with genders, it is the gender of a noun which determines its agreement pattern, not its declension type. For example, in Greek, nouns of the first declension are generally feminine and nouns of the second declension are generally masculine. However, the noun *mathetes* "student" is a first declension noun, taking first declension endings, but it is a masculine noun because it takes agreement with a masculine article and adjectives. In many languages the strong correlation between declension and gender admits many exceptions. Declension type should therefore not be confused with gender.

Given this, it is worth considering why the term “gender” is used at all.

Since so many languages make use of gender classifications which seem arbitrary and totally unrelated to natural sex, or correspond to natural sex only loosely, a more neutral term such as "noun class" or "agreement type" might be less confusing (Harris, 1991). Indeed, studies of Bantu languages, where it is not unusual to find as many as 16 genders, generally refer to noun classes rather than genders. Corbett (1991) cites two reasons the term gender continues to be used. The first is simple etymology: the word “gender” comes from Latin *genus* meaning "type". Secondly, and more importantly, the distinction between grammatical genders is ultimately rooted in semantic distinction. All languages which distinguish grammatical genders do so based on some semantic principles. Corbett and Fraser (2000) define a continuum of languages with “semantic assignment” at one end, progressing through “predominantly semantic assignment” to languages in which formal factors (either morphological or phonological) dictate the majority of gender assignment. Languages with “semantic assignment” would be those in which the grammatical system is transparently related to the semantic system. For example, Tamil (a Dravidian language) and Godoberi (a Caucasian language) have rigidly semantic systems in which male humans take masculine agreement, female humans take feminine agreement, and all other nouns take neuter agreement. Diyari, an Austronesian language of Australia, has two genders, one for all female entities (human, animal or mythical) and another for all other nouns, whether male, neuter, or indistinct, animate or inanimate. Toward the other end of the spectrum are languages like

Spanish and French, where the semantic basis for gender classification is miniscule compared to the “semantic residue” – nouns assigned to gender classes based on morphological and phonological similarity to semantically assigned nouns (or, possibly, based on arbitrary or unrecoverable semantic criteria). However, according to Corbett and Fraser (2000) no language is known to exist in which the formal system is not ultimately based in semantic distinctions. Such a hypothetical language would be one in which, for instance, male and female referents were assigned to different genders with equal frequency. Even in languages where gender is almost entirely discernable from formal factors (e.g., Qafar) there is “considerable overlap with semantic criteria” (Corbett, 1991; 307-8), and in the rare cases where semantics and form conflict, semantics trumps form in gender assignment (Corbett and Fraser, 2000). Therefore, despite the common lament of language learners that “gender is so arbitrary!” the semantic core of gender systems must be acknowledged.

2.2 What’s gender good for anyway?

Given that many languages do not have grammatical gender and seem to get along just fine, it is worth pursuing briefly what role gender fills in +gen languages.

One role that gender plays is a semantic role, differentiating between certain lexical items. For example, the kinship terms denoting male and female relations share the same stem and differ only in gender marking: e.g., *abuelo*

'grandfather' and *abuela* 'grandmother.' In another example, *manzana* 'apple' and *manzano* 'apple tree,' the use of gender may ultimately be related to the size difference between the fruit and the tree (Lucy, 2000). Lucy elaborates this role much more convincingly for classifiers than for gender, however, and it may be more likely that a +gen language makes use of the gender system it already has for this purpose, rather than this purpose fueling a gender system.

Another role of gender is aiding in disambiguating reference, or what Corbett (1991) calls "reference tracking." Spanish is a pro-drop language, which means that in many cases subject nouns may be omitted from a sentence, as long as the referent of the omitted noun can be discerned from context or from other elements of the sentence. By requiring modifiers to agree with the noun they modify, the referent of an omitted noun is more easily recoverable because the potential nouns to which the modifiers are semantically attached is narrowed to only nouns of the gender matching the modifiers. For example, the sentence *Dame por favor la roja* ("Please give me the (f) red [one]") can refer only to an object which is feminine in gender, because the definite article *la* and the adjective *roja* are both feminine, having taken feminine agreement marking from the deleted noun. Thus the red thing may be (for example) a chair *silla*, a skirt *falda*, or a suitcase *maleta*, because they are all feminine, but not a book *libro* or a glass *vaso* because these are masculine.

Psycholinguistic research indicates that comprehension of anaphoric pronouns may be facilitated by gender cues. In a self-paced reading and comprehension study, Carreiras, Garnham and Oakhill (1993) showed that

comprehension of Spanish sentences with an anaphoric pronoun (such as sentence 1c) was faster when preceded by a sentence with only one gender-correct referent (1a), than when preceded by a sentence with two gender correct referents (1b).

(1a) *La arena se pegó al helado*
The sand (f) stuck to the ice cream (m)

(1b) *El polvo se pegó al helado*
The dust (m) stuck to the ice cream (m)

(1c) *porque lo arrojaron al suelo*
because it (m) they threw on the ground
(because it was thrown on the ground)

That is, subjects were quicker at comprehending that the pronoun in (1c) referred to *helado* ('ice cream') when the item that 'stuck to' the ice cream was of a different gender, and therefore was not a possible referent for the pronoun.

Cacciari, Carreiras and Cionini (1997) further showed that gender agreement may facilitate language processing. Their experiment used epicenes and ungendered words in Italian, both of which can be used to refer to referents of either sex. Epicenes are words with morphosyntactic gender such as *la vittima* 'the victim' and always take the agreement of the formal gender class. Ungendered words, such as *l'erede* 'the heir,' take the gender agreement of the referent – that is, masculine if the referent is male and feminine if the referent is female. Subjects were faster at reading and comprehending sentences in which the semantic gender of a pronoun anaphor (*lei* 'she' or *lui* 'he') matched the formal gender of an epicene antecedent than they were when there was a gender

mismatch between the two, or when the antecedent was an ungendered word.

“These results suggest that the surface gender information incorporated in the antecedent, presumably at a lexical level, plays an important role in the on-line interpretation of anaphoric expressions (Cacciari, Carreiras and Cionini, 1997; 529.)

It is also possible that gender agreement serves as a redundancy feature in language. Agreement between a noun and its modifiers may serve to increase the signal:noise ratio, making it easier to understand what an interlocutor is saying in sub-optimal conditions.

2.3 The gender system of Spanish

The Spanish gender system is a straightforward one of agreement between nouns and their determiners and adjectives within DP, and between nouns and predicate adjectives. Masculine nouns always take masculine articles (definite *el* or indefinite *un* in the singular) whereas feminine nouns always⁴ take feminine articles (definite *la* or indefinite *una* in the singular). Adjectives, usually following the noun, are also inflected for gender (although there are some invariant adjectives which are used with both genders, such as *azul* ‘blue’). Adjective gender endings overwhelmingly correspond to noun gender endings; the masculine *-o* endings and feminine *-a* endings on nouns are the same as the masculine and feminine endings on the adjectives.

⁴ With some exceptions which shall be dealt with below.

- a) *el zapato rojo* 'the red shoe'
- b) *la camisa roja* 'the red shirt'

Of course, to some extent this is an oversimplification. Studied in depth, the phonological rules determining gender are complex and admit many exceptions, and some of the rules operate within small sets of words whose meanings aren't generally known by young children (Teschner and Russell, 1984). Therefore it would not be expected that learners even in a naturalistic setting such as the immersion program studied here would acquire all the correspondences, or even most of them, within the first few months of exposure to the language. Because of this, a lengthy discussion of all the Spanish gender rules is unnecessary, and the interested reader is referred to Butt and Benjamin (1988) and to Teschner and Russell's (1984) statistical analysis of Spanish gender. Instead, this discussion focuses on the aspects of the Spanish gender system which will be relevant to the experiment described in Chapter 4.

All nouns in Spanish are either masculine or feminine. In general, nouns denoting male humans and the males of some higher animal species are masculine nouns, while female humans and the females of some animal species are denoted by feminine nouns. There are many animals, however, whose names have unvarying gender (eg, *la jirafa* "giraffe", *el sapo* "frog", *el panda* "panda") regardless of the sex of the referent. Masculine is the unmarked gender, in the sense that when the sex of the referent is unknown or when the referent is a group of individuals of both sexes, the masculine form is used. Inanimate objects are either masculine or feminine.

Statistically speaking, the gender of a noun is largely predictable given the word ending. The clearest example of this is that nouns ending in *-o* are almost always (99.87%) masculine and nouns ending in *-a* are usually (96%) feminine (Teschner and Russell, 1984). These two endings account for more than 28,500⁵ of the 41,888 nouns in Teschner and Russell's (1984) analysis, or 68%. Nouns in these categories will be referred to as "canonical." In addition, nouns ending in *-d* are overwhelmingly feminine (97.57%) as are nouns ending in *-ión* (99.4%), and nouns ending in *-e* are largely masculine (89.35%), accounting for another 6248 words. This makes a total of approximately 34,800 words, or 83% of Spanish nouns which can be classified as "overwhelmingly predictable" based on morpho-phonological features.

There are exceptions to these generalizations, as in any language, and it is important to consider whether these exceptions may be frequent or salient enough in the K-1 classroom input to constitute a distraction to children's acquisition of the general "rules." For example, there are some very frequent exceptions to the *-o/-a* rule (e.g., *el día* "day", *el clima* "climate", *el poema* "poem", *el mapa* "map", *el tema* "subject", *el idioma* "language"; *la mano* "hand", *la radio* "radio", *la foto* "photo") which are likely to be commonly used in the classroom. Furthermore, while *-e* is overall predictive of masculine gender, Teschner and Russell (1984) report that more than a quarter of the *-e* nouns in a dictionary of the 5,000 most frequent Spanish words are feminine (e.g., *la calle* "street", *la noche* "night", *la tarde* "afternoon", *la gente* "people", *la clase* "class" are all in the list of the top

⁵ An exact count of the over 16,000 *-a* nouns is not available (Teschner and Russell, 1984).

500 Spanish words). Thus statistical tendencies toward predictability may not accurately represent the subset of words actually being used and acquired in the kindergarten and first-grade classrooms.

Masculine nouns co-occur with masculine determiners (e.g., *el, un*), and with adjectives and passive participles marked for masculine gender by the suffix *-o*. In most cases feminine nouns co-occur with the feminine determiners (*la, una*), and with adjectives and passive participles marked feminine by the suffix *-a*. The correlation between these agreement endings and the large number of *-o/-a* nouns may be a help to children in acquiring the basic grammatical gender distinction. However, even here there are exceptions. Words beginning in a stressed *a* –or *ha*– undergo a postlexical phonological rule causing them to take masculine articles in the singular (but not the plural). Butt and Benjamin (1988) list several examples not unlikely to be used in a young children’s classroom (e.g., *el agua* “water”, *el águila* “eagle”, *el Africa* “Africa”, *el área* “area”, *el habla* “speech”, *el hambre* “hunger”, *el hada* “fairy”).

2.4 The gender system of English.

English has only a very limited gender system. Inherent gender is based entirely on natural gender, and is encoded only on the 3rd person singular pronouns in the nominative (he/she/it), accusative (him/her/it), genitive (his/hers/its), and reflexive cases (himself/herself/itself). It is also sometimes inherent in specific nouns, usually describing females (waitress, actress, stewardess) but many of these nouns are undergoing a cultural change to more

“gender neutral” terms (server, actor, flight attendant) and are in less common use than they were 20 years ago. Most importantly, English has no grammatical agreement between nouns or pronouns and their associated determiners and adjectives; that is, English has no derived gender. This lack of grammatical gender agreement is what qualifies English as a –gen language.

Spanish pronouns are similar to English. The nominative pronouns *él* and *ella* correspond to ‘he’ and ‘she’ respectively, and the accusative *lo* and *la* to ‘him’ and ‘her’.⁶ However, where English uses the gender-neutral ‘they’ in the plural, Spanish also distinguishes between a group of females *ellas*, and a group of males *ellos*. However, if the group referred to is of mixed sexes, the pronoun used is the masculine plural.

In view of this typological difference between Spanish and English, what predictions are we able to make about how English L1 learners of Spanish will approach the task of acquiring Spanish gender? One possibility is that learners will initially hypothesize that there is no gender feature in Spanish. That is, they will transfer their L1 feature inventory, which does not contain a grammatical gender feature, to their L2. They will also transfer their pronoun gender categories of ‘he’ referring to males and ‘she’ referring to females. However, given positive evidence for the existence of gender in Spanish, learners will incorporate gender into their interlanguage. Since the +gen evidence they encounter will coincide in cases where natural gender is present with their own L1

⁶ Not all languages distinguish natural sex in spoken language. Mandarin, for instance, uses *ta* (sing.) and *tamen* (plural) for humans of both sexes, although there is a distinction in written Mandarin.

categories for natural gender, it is possible that they may hypothesize that natural gender feature as the basis for the grammatical gender feature in L2.

2.5 Psychological correlates of linguistic gender

Given that grammatical gender is related, if sometimes only loosely, to natural gender, it is relevant to ask what that relationship is. It is especially relevant here, in light of the fact that the children in this study will be moving from an impoverished gender system which is entirely transparent to natural sex toward an arbitrary morpho-phonological gender system which overlaps in a small part with the L1 system. In addition, it is reasonable to question whether the semantic basis of grammatical gender in a language is in any way relevant to speakers of the language. To what extent are the concepts of natural gender and grammatical gender intermingled for individual speakers of the language? Does the encoding of grammatical gender in a language, and the form and pervasiveness of this encoding, affect children's attainment of gender identity (that is, recognition of themselves as either a boy or a girl)? Does the presence or absence of grammatical gender affect the way in which children and adults classify objects as having masculine or feminine qualities? Does the fact that grammatical gender has a semantic basis aid children in acquiring the grammatical gender system?

A few studies have addressed such questions. Guiora et al. (1983) find a direct relationship between the pervasiveness of gender agreement in a language and the age at which children achieve gender identity (i.e., understanding of their

own status as a “boy” or “girl”). They tested children from three languages, Hebrew (with pervasive gender agreement), English (with minimal gender agreement) and Finnish (with zero gender), on a standardized test of gender identity attainment, and found that the age at which children correctly associate themselves as either male or female correlates largely with the amount of gender loading in the language (i.e., Hebrew-speaking children attained gender identity earliest, Finnish latest). Mills (1986) noticed that children speaking a +gen language (German) made fewer mistakes in pronoun reference than English speakers. Both Mills’ and Guiora et al.’s findings are consistent with Mill’s suggestion that languages where gender-marked pronouns are extended to all nouns rather than just humans afford more practice with the pronominal system, facilitating earlier mastery.⁷

Clarke et al. (1981) addressed the question of whether adults from languages with different levels of gender encoding perceive inanimate objects differently in terms of gender. Arabic speakers and English speakers were shown pictures of different inanimate objects and asked to give their “first impression” as to whether the word denoted an object which was conceptually more female or male. Arabic speakers showed a marked tendency to answer in accordance with grammatical gender. Interestingly, this contrasts with Guiora and Acton (1979) who tested English and Hebrew speakers in the same task (Hebrew, like Arabic, having two genders and pervasive agreement marking). In that study, no significant difference between English speakers and Hebrew speakers was

⁷ Mills’ (1986) study is discussed in greater detail in Chapter 3.

detected, suggesting that all adults, regardless of language, categorize objects in the world using similar criteria unrelated to gender marking in the language.

Such studies are fascinating but they are also fraught with problems, a fact which does not escape the notice of the authors. First of all, it is very difficult to tease apart the influence of language from the influence of culture. Language is an element of culture, and cultural values are often evident in language (e.g., through idioms, etc.). Therefore, in investigating how *language* affects perceptions of gender, investigators must take great pains to filter other gender-oriented cultural aspects out of the study. As Clarke et al. (1981) note, when language is the independent variable, separating language from culture is essential to the veracity of the study. Therefore any study which attempts to pin a particular effect on language bears the burden of proof to show that it is indeed an effect of language and not, even in part, an effect of some other aspect of the relevant culture or society.

To my knowledge only two studies have attempted this. In a smaller study also reported in Guiora et al (1983), Swedish-speakers in Finland were compared with Finnish-speakers in Finland (Swedish, like English, has minimal gender) on the same standardized measure of gender identity used in the larger study with Hebrew- and English-speakers. In addition, a cultural survey was given to both Swedish-speaking and Finnish-speaking Finns to determine the extent to which male and female gender roles were important in their culture. Swedish-speakers reported themselves to be much more gender-neutral in their expectations for their children while Finnish-speakers considered themselves very traditional in their

expectation of their children's gender roles. Nevertheless, the Swedish-speaking children showed an earlier attainment of gender identity than the Finnish-speakers, leading the authors to suggest that the effect under view is really an effect of language rather than of other cultural factors.

A second, more rigorous study which attempts to extract purely linguistic information from cultural influences is by Sera, Berge and Pintado (1994). They conducted three experiments with Spanish- and English-speakers in which they showed subjects pictures of objects, half of which were grammatically masculine in Spanish, the other half of which were feminine. Half of the subjects saw the pictures only, the other half of the subjects saw the pictures and heard the experimenter label the picture. In each case, the subjects were asked to indicate whether they considered the object to be more male-like or more female-like. Spanish-speaking subjects from 2nd grade to adult were significantly more likely to classify objects in accord with their grammatical gender when pictures were labeled than when they were not labeled. This suggests that there is an effect of grammatical gender independent of what the authors term "folk-theories" of gender which may be culturally defined.

Furthermore, the biases of the researchers cannot be discounted. For example, Mullen (1990) notes that English-speaking children tend to associate natural objects with feminine characteristics and manufactured objects with masculine characteristics, a tendency also noted by Sera, Berge and Pintado (1994). It is not impossible that such tendencies are exhibited in the results of

studies like those reported here, but pass unnoticed because they are not the focus of observation or are masked by other factors.⁸

Martinez and Shatz (1996) showed differences in how monolingual Spanish- and English-speaking preschoolers categorize pictures according both natural and grammatical gender. Children were shown drawings of humans and known objects. Of the inanimate objects, half were grammatically masculine and half were grammatically feminine. First the children were asked to sort them in any way they chose. Next, the experimenters asked the children to sort by natural gender. Finally, Spanish-speaking children were asked to sort pictures based on grammatical gender. During the free sort, four of the English-speaking children used a natural gender distinction, while only one Spanish-speaking child sorted according to natural gender. On the other hand, six Spanish-speaking children sorted according to grammatical gender in the free sort. When asked to sort by natural gender most of the English-speaking children completed the task successfully. However, of the Spanish speaking children only slightly more than half completed the natural gender sort successfully. Three sorted instead by grammatical gender and five showed no pattern. The authors suggest that this shows the strength of formal linguistic properties over semantic gender information. They also note the fact that more Spanish-speaking than English-speaking children failed to sort using any pattern in this task suggests that the use

⁸ Mullen's (1990) findings do not bear directly on Clarke et al's (1981) study, as they kept the number of natural and manufactured objects constant in each group of words they presented. However, the caution is not unwarranted.

of the terms “masculine” and “feminine” to describe both natural and grammatical gender may have caused confusion.

Whether the potential psychological correlates of gender have any influence on children’s L2 acquisition is an open question. Chapter 3 will review several studies which show that children learning their L1 are oblivious to any correspondence between natural sex and grammatical gender. This is, of course, to be expected, as L1 learners are too young to be aware of sexual differences much beyond the fact of there being two sexes and possibly their parents’ membership in one or the other category. The children in this study, however, having reached an age where they are aware of sexual differences and are beginning to build their own identity as a member of either sex, may be more attuned to the correspondences than L1 learners are, and therefore may utilize their understanding of the correspondence in their L2 acquisition of grammatical gender. First, we examine the theoretical underpinnings of L1 acquisition, which will set the context for a review of the L1 gender literature.

2.6 Conclusion

Chapter 2 has explained what grammatical gender is and the purpose it performs in language, as well as how grammatical gender is manifested in both Spanish and English. Furthermore, it has illustrated that there may be psychological correlates of grammatical gender which influence perception, and vice versa. Chapter 3 will illustrate how gender fits into an overall theory of L1 acquisition, and review some of the literature most relevant to the present study.

Chapter 3

To understand whether children acquire grammatical gender in the L2 in the same way they acquire gender in their L1, we must understand what we know about L1 acquisition. This chapter will review research on L1 gender acquisition, placing it in the context of what will be called the “syntax first” theory of Maratsos and Chalkley (1980). The chapter will briefly outline the competing hypotheses of L1 acquisition, then review this early work, and then review some examples of gender L1 research which support the “syntax first” position.

3.1 Acquisition of gender in L1

One important question (of many in L1 research) is how children develop grammatical categories. Children must form word classes such as noun, verb, adjective, and preposition as part of their emerging grammatical competence. How do children go about the business of discerning verbs from adjectives, verbs from nouns, nouns from pronouns from prepositions from determiners? The reason this is important is that grammatical gender sets up distinctions between classes of nouns. How does a child learn the grammatical concepts of masculine, feminine and neuter (or, whatever the gender categories for his language are)? It is reasonable to suggest that whatever method a child has for discerning the larger categories of grammatical word classes would be the same method he uses for dividing those word classes into subclasses, such as gender subclasses.

An early but influential idea was that the necessary grammatical categories were innate (Chomsky, 1965). There was no need to formulate a theory of how children arrived at a set of grammatical categories, because categories like noun, verb and adjective were concepts the child brought to the task of acquisition, concepts which were fully formed and tailor-made for language. This position grew out of the theory of transformational grammar, which held that, because the elements important to category formation were not always apparent in the surface structure of the language input to which the child was exposed, the categories therefore could not be “learned” by attention to the input. However, as grammatical theories have changed, so have theories of language acquisition. Later research has proposed that children must construct grammatical categories for themselves and if they bring anything innate to the task, it is an innate predisposition to attend to the input in such a way that they are able to discern the relevant linguistic information allowing them to construct those categories. For our purposes, a nativist position would be entirely unhelpful in the question of how gender categories are distinguished. Since gender categories vary from language to language and are altogether absent in many, it is therefore most instructive to compare the other two major hypotheses of grammatical gender construction.

A second hypothesis which attempts to explain the child’s grammatical category development might be termed the “semantics first” hypothesis (Schlesinger, 1971; 1988; Slobin, 1966). Schlesinger articulates a theory of “semantic assimilation” using the semantic roles that have been shown to be

present in children's earliest two-word utterances: agent, action, location, possession, patient. He concentrates primarily on the agent-action relationship. The argument proceeds along these lines. Children see the world in non-linguistic terms first. They see and note that people (and other beings and objects) exist. They note that people engage in activities such as running, talking, eating, sitting, and sleeping. They interpret their surroundings in terms of who they see doing what, and hear language which encodes those relationships. In seeking language to encode relationships, a child begins to form linguistic categories based on those semantic relationships, such as "agent" and "action". Clearly not all subjects a child hears will be agentive, and not all verbs will be actions; the subject and verb in the sentence "The doggie likes the bone" are neither. But the narrow relational categories that are formed initially will eventually expand to admit meanings that are similar, (e.g., agent → subject, action → verb) leading in time to a changeover where semantics is no longer primary. At this point relational categories may expand based on either semantic or syntactic criteria.

These theories have many adherents and much important work has been done over the years which continues to fuel the debate. The theory which best explains the relevant research on gender, however, is the "syntax first" hypothesis, expressed most clearly by Maratsos and Chalkley (1980; see also Braine, 1987; Maratsos, 1988). Therefore, a bit more attention will now be devoted to this theory and placing gender research within it.

3.2 The “syntax first” approach

According to Maratsos and Chalkley (1980), one of children’s primary linguistic tasks is to categorize words into form classes or categories similar to the grammatical categories in adult language, noun (N), verb (V), adjective (Adj) and so on. Since these are essentially syntactic categories that the child must construct, it is logical to begin with the hypothesis that children use syntactic criteria for constructing them. This is exactly what Maratsos and Chalkley believe they do. Children build grammatical categories by paying attention to the syntactic and morphological frames in which words appear in the input. Maratsos and Chalkley illustrate how the child accomplishes this task using the categories of verb and adjective as examples. Both verbs and adjectives are words which denote a relation between entities. The verb *like*, for example, denotes a relationship of liking between the entities, one who likes something and one which is liked. The adjective *nice* denotes a relationship between an entity and the quality of being nice. Maratsos and Chalkley suggest that the child assigns words to categories on the basis of the distributional patterns in which the words appear in the input.

For example, imagine the following set of sentences as initial input.

- (1a) You liked the banana.
- (b) Joey likes the ball.
- (c) Mommy talked to the man, didn’t she.
- (d) Joey talks to the bears.

Given sentences like those in (1), a child can begin to see a correspondence between words that take –s endings in a present tense situation, and words that take –ed in past tense situations. He associates the words *like* and *talk* through their common occurrence in identical syntactic frames. In this way he begins to build a network of morphological and syntactic frames that define categories, and in turn, adds new words to his vocabulary in the correct category. If the child subsequently hears a sentence like (2a), he can, through his category-building mechanisms, assume that a sentence like (2b) is also possible.

- (2a) The puppy walks on four legs.
- (b) The puppy walked yesterday.

Imagine that the child next notices sentences like those in (3)

- (3a) Joey doesn't like bananas.
- (b) Mommy didn't talk to anyone.

Given such input, the child can hypothesize that the forms *doesn't talk*, *doesn't walk*, *didn't like*, and *didn't walk* are also acceptable in his language, because as members of the same grammatical category, all those words are admissible in the same frames. Maratsos and Chalkley convey this idea of building networks with a diagram similar to figure 3.1.

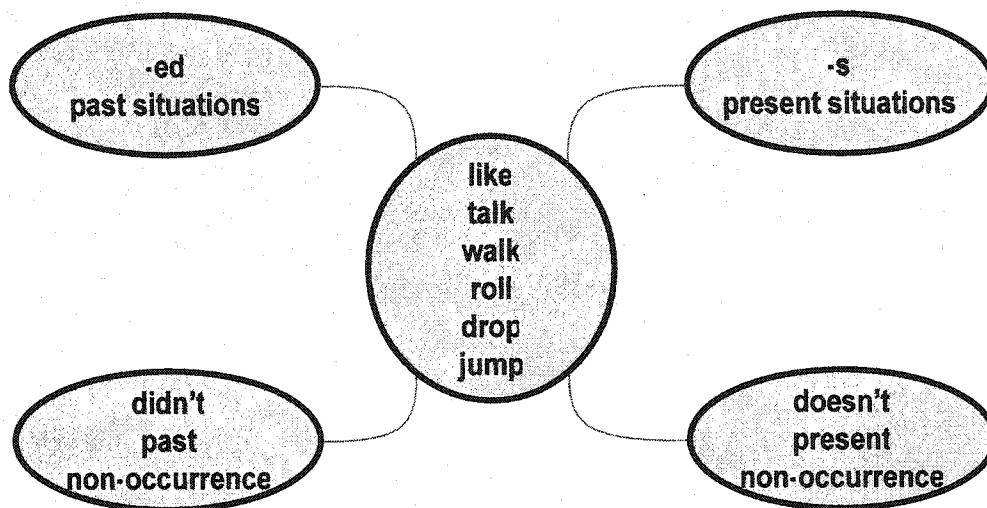


Figure 3.1. Network of verb correspondences

In a similar way, a child might notice correspondences between the morphological and syntactic frames in which words such as *noisy*, *pretty*, *happy*, and *naughty* appear. For example, a set of input sentences such as those in (5) would lead the child to build a network of correspondences between those words as diagrammed in Figure 3.2.

- (5a) The big boy is noisy!
- (b) The dog was noisy.
- (c) This is a pretty dress!
- (d) This isn't a pretty doll, is it.
- (e) The little girl is happy.
- (f) The little girl wasn't happy yesterday.
- (g) That is a naughty thing to do!
- (h) The boy in the video was naughty.

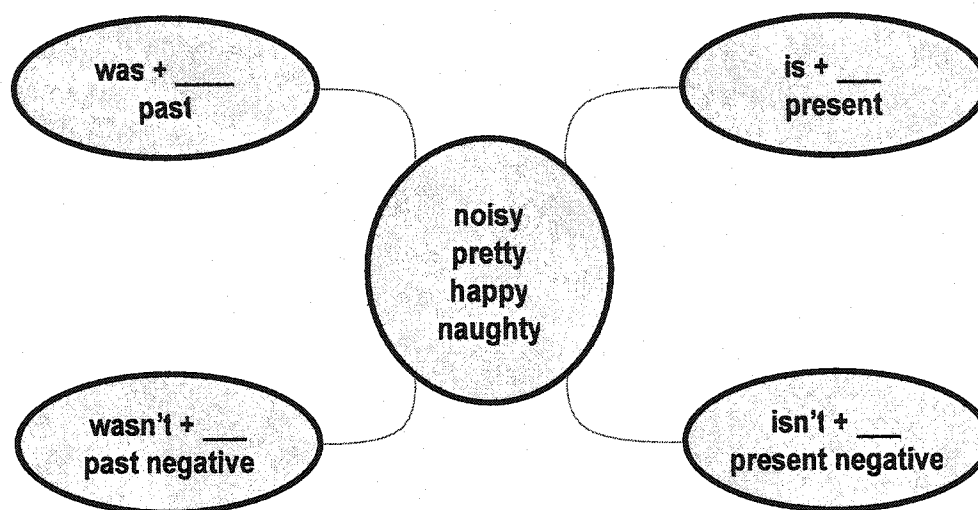


Figure 3.2. Network of adjective correspondences

The important point to note here is that there is no way the child could form these networks of syntactic correspondences through semantic analysis alone. A purely semantic analysis would be one in which words were assigned to their grammatical category on the basis of meaning. For example, since both adjectives and verbs are predicates, a child using a purely semantic analysis could logically but erroneously assign them to a single class of “predicates.” On the other hand, it could be argued that a child’s semantic analysis might be deeper than “predicate,” for example, he might hypothetically distinguish the two categories by assigning all words that seem to denote actions to one class (V) and all words denoting states of being to another class (Adj). While it is true that verbs tend to denote actions and adjectives tend to denote states, the correspondence is far from perfect, as the words in the examples above show.

Like is a mental state whereas *talk* is an action, and yet both fit into the verb category by virtue of the identical syntactic and morphological frames in which both occur. Similarly, *happy* and *pretty* denote states as we expect adjectives to do, but *naughty* and *noisy* presuppose some activity more than state. Thus the child would have no way to distinguish between categories based solely on meaning, or if he did try to use such criteria, he would ultimately come up with entirely the wrong generalizations.

Of course, the networks that the child constructs are not merely correspondences between morpho-syntactic frames in which words of particular grammatical categories occur. While children are constructing categories and networks, they are also learning the meanings of the words that fall into these categories and fit into the frames. Thus the child is constantly building both semantic and syntactic understanding of the language. The child's syntactic analysis cannot be entirely independent of semantic content. The networks themselves include information about the meaning of the particular morpho-syntactic frame. For example, it is not merely the case that words which take -s endings can also take -ed endings. Rather, words which take -s endings *in the context of a present action* also take -ed endings *in the context of a past action*.

Furthermore, it is certainly true that some semantic distinctions are relevant, even necessary, to discerning correct syntactic patterns. For example, the progressive morpheme -*ing* attaches to some verbs to show a progressive action, such as *talking* and *running*. It also attaches to the *be* verb preceding certain adjectives with more active meanings, such as *noisy* and *silly* to suggest a

purposeful behavior, such as *He is being noisy* or *he is being silly*. The same verbs and adjectives which take *-ing* are those which can also be used in the imperative, e.g., *Run to the door!* *Be silly!* Verbs which denote mental states such as *know* and *want*, and prototypical adjectives denoting states such as *red* or *hungry* do not participate in either of these syntactic frames. Only a semantic analysis of the meanings of the verbs and adjectives which do and do not fit these morpho-syntactic frames could lead a child to discern this.

Theoretically there is no reason that the child *couldn't* begin with semantic criteria. One argument might be that children begin the task of classifying words using purely semantic criteria and then, when the semantic analysis breaks down, switch to a syntactic analysis. However, if this were the case, we would expect children to make mistakes in the classifications that they give words, and Maratsos and Chalkley note that this type of mistake is very rare in the literature. Errors such as "she happied" are virtually unattested.

If children are in fact using a morpho-syntactic type of analysis to create categories, it should be the case that the major categories (N, V, etc) are formed fairly early, as hearing only a small number of words in two or three common morpho-syntactic frames would allow the child to begin forming these classes. Valian (1986) has shown that children as young as 2 years old have indeed already created some grammatical categories in their early grammar. She examined tape recorded data from six children ages 2 to 2;5, with mean length of

utterance⁹ between 2.93 to 4.14, for evidence of the existence of the classes Det, Adj, N, NP, P, and PP by looking for distributional regularities. For example, in English, adjectives but not determiners can be used serially. If the child's language conforms to this pattern, this is evidence that the child's grammar includes a distinction between Det and Adj. That is, the child has acquired these syntactic categories. Valian found that all of the children had acquired the categories Det, N, NP and P, and all but the youngest child showed evidence of the categories Adj and PP. Valian's results show that children are clearly sensitive to syntactic regularities by a very early age, supporting Maratsos and Chalkley's theory of syntactic analysis for the creation of grammatical categories.

Remember that grammatical gender is itself a syntactic phenomenon. Although there are clear correlations between the natural sex of a referent (if it has such) and its gender, these correlations often break down (as in Greek *koritsi* or German *mädchen*, "girl" which are grammatically neuter) and for the majority of nouns no such correlation exists. Thus grammatical gender is an excellent example of a set of categories which may be figured out primarily through syntactic data. Of course it is possible that children might begin with semantic information (e.g., my mommy and sister are feminine, my daddy and brother are masculine, etc.). But children's understanding of biological categories such as sex is not developed at this point. And of course, the semantic line of analysis would fail the child almost immediately. Given that the children can already be

⁹ Mean length of utterance (MLU) is a measure of children's syntactic development, calculated by counting the morphemes in a series of utterances and taking the mean number of morphemes per utterance.

shown to be capable of a syntactic analysis to create other sets of categories (Valian, 1986), it is far more logical to assume that the same sort of analysis will be the child's default analysis regarding grammatical gender. Accordingly, the following section will detail some of the most important studies on children's native acquisition of their language's gender system, which support Maratsos and Chalkley's (1980) theory regarding syntactic analysis.

3.3 Gender acquisition research supporting "syntax first"

Most of the research in this area regards first language acquisition of languages with grammatical gender systems. Gender systems differ greatly from language to language, and the extent to which natural gender (a semantic feature), inherent gender (generally a morpho-phonological feature) or derived gender (a syntactic feature) are clearly apparent in the language may greatly impact their salience to learners. It is therefore necessary to examine evidence from as many languages as possible to determine which gender features are most salient to children. The following section will detail some of the notable research with children learning a +gen language as a native language. We will also look in greater detail at the acquisition of the Spanish and English gender systems, since these are the systems relevant to the present research.

3.3.1 French

The seminal research on gender acquisition in French was done by Karmiloff-Smith (1979). She conducted five experiments with children aged 3 to 11, investigating the importance of semantic, syntactic, and phonological cues to gender in French. French nouns are all either masculine or feminine. Some noun endings are good predictors of gender (e.g., nouns ending in *-ois*, *-on*, and *-eau* tend to be masculine whereas nouns ending in *-elle*, *-ienne*, and *-ette* tend to be feminine) while other endings are more ambiguous (e.g., *-ile*, *-ique*, and *-ire*). Gender is also marked on determiners, adjectives, pronouns, etc., which could provide a clue to noun gender. Karmiloff-Smith found that young Francophone children relied primarily on phonological information in the noun ending to determine gender, rather than on syntactic or semantic information.

In her experiments, Karmiloff-Smith showed the children pictures of imaginary people who were obviously male or female, as well as pictures of imaginary animals and objects. Each drawing was labeled verbally with a French-sounding nonsense word whose ending was either typically masculine or feminine, or one which was typically ambiguous. All five experiments used the same set of nonsense words and the same set of drawings. Pictures were introduced always without definite articles, so the definite article used by the child when subsequently referring to the picture was taken as the indication of

which gender the child considered the noun to be. The general procedure for Experiments 7-10¹⁰ was as follows (in French):

Experimenter shows child picture of an imaginary object, animal, or person.

Exp.: "Here is a picture of a plichette. What is it?"

Ch.: "A plichette."

Exp. shows a nearly identical picture.

Exp.: "Here's another picture. What do you think it is?"

Ch.: "Another plichette, but a grey one."

Exp. hides one picture, or places an object on one of them, or performs another action.

Exp.: "What have I done?"

Ch.: "You have covered up the green plichette."

In experiment 7, pictures were introduced using an indefinite article the gender of which corresponded with the phonological word ending, if present, providing a syntactic cue to gender. Such cases where two or more cues to gender both point to the same gender will be termed "converging cues." In experiment 8, pictures showed two entities and were introduced using the quantifier *deux* "two" which does not show agreement, thus no syntactic cue was provided. In experiment 9, object and animal pictures were presented with indefinite articles which did not agree with the phonological ending of the noun (e.g., *une bicron*). Such cases where two or more cues to gender point to different genders (e.g.,

¹⁰ Because the gender experiments are part of Karmiloff-Smith's much larger study on determiners and reference generally, the experiment numbers are 7-11, rather than beginning with '1'. Experiment 11 was slightly different and it will be discussed separately.

feminine articles co-occurring with masculine endings) will be termed “conflicting cues.” In experiment 10, noun ending and natural sex were in conflict, and the quantifier *deux* was again used so no syntactic cue was available. Table 3.1 shows the cues being investigated in each of the first four experiments. Experiment 11 was slightly different. In this experiment, Karmiloff-Smith showed the children a male picture with a male name and asked the children to invent a name for their female counterparts “from the same country.” The results of Experiment 11 will be considered separately.

Table 3.1: Cues under investigation in Karmiloff-Smith (1979) Experiments 7-10

	Linguistic Cue	Semantic Cue (when present)
Experiment 7	Article and noun ending converging	Converging
Experiment 8	Noun ending only	Converging or not present?
Experiment 9	Article and noun ending in conflict	Not present
Experiment 10	Noun ending only	Conflicting

Experiment 7 offered three different combinations of cues, which will be referred to as Conditions A, B, and C. The combinations in Experiment 7 were as follows:

Condition A: all three cues (semantic, phonological, and syntactic) converging;

Condition B: phonological and syntactic cues only, converging;

Condition C: syntactic cues only.

When the phonological cue was present (Conditions A and B), children's choice of definite article agreed (89-100%) with the masculine or feminine phonological cue¹¹, but when the noun ending was ambiguous (Condition C) scores dropped (64-91%) despite the presence of the syntactic cue. There are several possible interpretations of these data. One possibility is that phonological information is primary for children and its absence hinders children in their assignment of gender. Alternatively, the confluence of two linguistic cues aids children in their assignment of gender and the absence of one or the other cue hinders this assignment. On the other hand it may be that *any* two converging cues (e.g., semantic and syntactic, or semantic and phonological) might aid children in their assignment of gender.

The data from Experiment 8 are instructive in deciding in favor of the first interpretation; that is, that phonological information is the most important cue for children. In Experiment 8 children had only the phonological cue available (presumably without natural gender cues, although this is not specified). The data seem to show that children are more reliant on the phonological (noun ending) cue than on the syntactic (article) cue. However, the data also seem to suggest

¹¹ Interestingly, 3- and 4-year-olds did slightly better in Condition B with no natural gender cue whereas 6-year-olds did slightly better with the natural gender cue. 5- and 7-year-olds showed no difference between the two conditions and the data for 9- to 11-year-olds is not reported.)

that the ability to use the phonological information depends on the exact nature of that information. The masculine ending *-rons* elicited nearly perfect masculine articles while the masculine ending *-iers* elicited masculine articles between 71-100%. Results were even lower for feminine endings. The *-ettes* ending elicited as little as 69% feminine articles, and the *-ines* ending as low as 57%.

Interpreting the results of the two experiments together is difficult since only a portion of the data is reported for each experiment and the parts which are reported are not entirely congruent. However, the two sets of data appear to reveal two things. First of all, given only *one* linguistic cue, children can make more reliable use of phonological information at the end of the noun than they can of syntactic agreement information in the article, *provided that* the information is considered reliable by the child. As noted in Chapter 2, in Spanish certain noun endings are highly reliable predictors of gender whereas others, despite being generally typical of one gender, are less reliable because of the number or frequency of exception. Karmiloff-Smith does not mention whether certain French noun endings are more typical examples of masculine and feminine gender than other noun endings, but data from Experiment 8 suggest that this is so.

Secondly, there appears to be a difference between children's ability to use masculine and feminine cues. Given an isolated syntactic cue, 4- to 6-year-olds more reliably used feminine articles than they did masculine articles (in 3-year-olds this trend was reversed; 7-year-olds were equally reliable for both genders, and no data were reported for older children). However, given an isolated phonological cue, most age groups (with the exception of 5-year-olds) more

reliably used masculine noun endings than feminine ones. 9-year-olds showed a strong tendency to treat all nouns as masculine regardless of noun ending, suggesting that older children may use masculine as a sort of default gender. This difference between masculine and feminine with regard to the two cues may indicate that for French speakers feminine is the “marked” value only syntactically. Hearing a feminine syntactic cue may trigger awareness of the marked gender value and lead to feminine gender assignment, whereas hearing a feminine phonological cue may be a less salient indicator of markedness.

Experiments 9 and 10 address the issue of what children do when they meet conflicting cues. In Experiment 9 the children were introduced to an object using an indefinite article of one gender paired with a noun ending typical of the opposite gender. Children under 5 tended to assign gender according to the type of noun ending rather than according to the gender of the indefinite article – that is, they used phonological rather than syntactic cues to gender. At 5 years old there is a dramatic increase (a jump from 37-46% up to 93%) in the assignment of feminine gender to nouns introduced with feminine indefinite articles, and 6- to 11-year-olds were more reliable in their ability to use syntactic cues of both genders (71-96%), but not nearly as reliable in cases of syntactic and phonemic conflict as they were in cases of convergence (84-100%). This suggests that even when children have acquired a level of gender understanding which takes greater note of syntactic cues, the phonemic cues still exert a strong force. However it is also interesting that when children begin to pay more attention to syntactic cues it is first noticeable with the feminine articles. This accords with the interpretation

of Experiment 8 above; when syntactic cues are attended to, the tendency is to make first use of the marked value, the feminine.

Experiment 10 examines conflict between phonological noun ending and natural gender. Until the age of 10, children assigned gender according to the noun ending rather than according to natural gender (only 13-34% reliance on natural gender). 9-year-olds used natural gender 46-48%, or roughly at levels according to chance. At age 10-11 there was an increase in the use of the masculine article for male pictures with feminine noun endings, however the same was not true for female pictures with masculine noun endings. These were reliably marked as masculine by children of all age groups. This again suggests that children may be using the masculine gender as a sort of default gender. A comparison of the results when there was no conflict (i.e., male and female pictures with ambiguous noun endings) confirms this analysis. Children produced reliably masculine articles when the picture was male but did *not* produce reliably feminine articles when the picture was female.

Experiment 11 takes a different approach, exploring children's understanding of the relationship between natural gender and the linguistic forms which encode it. In this experiment children were shown a male or female picture labeled with a masculine sounding nonce word, and asked to make up a word for a similar being of the opposite sex "from the same country." In this experiment, the child might choose to show the opposite sex by using the opposite indefinite article, or by changing the noun ending to one more typical of the opposite sex, or by a periphrastic circumlocution such as "a girl X" or "a female X"; and of course,

children might use a combination of these strategies. Children under 5 seemed to have difficulty with this task, as over 50% of them in all cases (and usually over 70%) either made no change at all or else refused to give any answer. For children 5 and over, Karmiloff-Smith's data seem to show that when the noun ending is ambiguous, children's primary strategy is to change only the article to the opposite gender. The older the children got, the more likely they were to also change the noun suffix. When given a female picture with a feminine noun and asked to create a noun for the male counterpart, children's primary strategy was to change both article and suffix. However, when given a male picture with a masculine noun and asked to create a name for the female counterpart, the development of a different strategy over time is most obvious. 5-year-olds were somewhat more likely to change only the article rather than both article and suffix (36-32%), 6-year-olds were much more likely to change only the article (57-32%) and 7-year-olds were somewhat more likely to change both article and suffix (45-49%), while 8-9-year-olds were much more likely to change both (up to 83% for 8-year-olds, 100% change of both for 9-year-olds.) Overall, children 6 and older were more likely to change both the suffix and the article, suggesting that children are less attuned to the gender function of the article itself than they are to the gender function of the noun suffix, or at least to the full complement of gender information. If children were really paying most attention to the noun suffix we might expect that some children would try to change only the gender of the noun suffix without changing the article as well. There is no evidence that the children ever did this, although this is never explicitly addressed.

Taylor-Brown (1984) notes that the typical masculine/feminine endings used by Karmiloff-Smith can be found on real French nouns which are sex-opposite pairs.¹² She suggests that some of the older children may have been aware of the natural gender significance of such endings. Thus if words with these endings were given for pictures of either inanimate objects or animate objects whose sex contradicted the gender ending, children might be inclined to dismiss the phonological information as irrelevant. In this case, the low percentage of older children answering in accord with the phonological ending would not be due to their non-use of phonological cues, but rather to their consideration and discounting of the information. This might also help explain the results of Experiment 8, in which certain noun endings evoked more consistently gender-biased responses than other endings. The correspondence between semantic gender and grammatical gender may be statistically more reliable for some endings than for others.

Karmiloff-Smith's gender study was part of a much larger study on determiners and reference, of which gender was a tangential part. Because of this, Karmiloff-Smith does not always give as much data as one would like for comparison. For example, in Experiments 7 Karmiloff-Smith tested children's gender attribution for all three cues in convergence, for phonological and syntactic cues in convergence, and for syntactic cue only. In Experiment 8 we are told that children had no syntactic cue but we are left ignorant of whether a

¹² Taylor-Brown (1984; 53) specifies the following pairs: *baron-baronne*, *chateau-chamelle*, *chien-chienne*, *fermier-fermière*, *cadet-cadette*, *chinois-chinoise*, and *français-française*.

semantic cue was available, or if all the pictures used here rather were inanimate objects. If all were inanimate objects, then Experiment 8 gives evidence about what children do with a phonological cue alone (which is what we assume based on repeated readings of the result), but if a gender cue was furnished, then this would constitute evidence for what children do when two cues are available, a combination of cues which does not appear to have been tested elsewhere in the project. In Experiment 9 we are given no information on what children do when they hear a gender-ambiguous suffix with a gender-specific indefinite article. One would imagine that they are likely to pay more attention to the article in such instances, but no data are given. In none of the experiments does Karmiloff-Smith appear to test children's reactions to conflict between syntactic and phonological cues when a semantic cue is available to possibly tip the balance in one direction or the other. Finally, Karmiloff-Smith's results are difficult to evaluate for significance because she gives only percentages of children's responses conforming to the predicted response. There is no indication of what percentage values she takes as a significant result, except that percentages around 50% seem to be considered as "chance". No statistical tests appear to have been conducted, so comparison between the percentages of the various age groups or between the different cues provided may seem to yield interesting results although in fact we do not know if the results reach statistical significance.

Karmiloff-Smith's work signals one of the most complete early attempts to discover what cues children attend to in acquiring their native gender system. However, it is interesting that the percentages of response that she reports rarely

show any change that can be attributed to development. That is to say, the percentages of predicted or unpredicted responses rarely seem to increase or decrease according to age. Graphs of the scores reported would show many peaks and dips throughout the 8-year age span. Other studies, however, have shown evidence that children's understanding of gender, and the strategies that they use to determine gender, do show developmental stages. It is to such a study that we turn now.

3.3.2 Hebrew

Levy's (1983a) study of Hebrew acquisition focuses on the cues children use to determine plural forms, which are in turn determined by the gender of the noun. Like Karmiloff-Smith (1979), Levy found that children rely primarily on morpho-phonological information rather than syntactic or semantic cues. In general, feminine nouns in Hebrew end in stressed *-a* or *-Vt* and take the plural ending *-ot*, all other nouns are masculine and take the plural ending *-im*.¹³ There are lexical exceptions to these rules which must be learned individually. Levy conducted two studies, one longitudinal and one cross-sectional, to determine what cues Hebrew children use to assign plural forms. In the longitudinal study, which followed her son from age 1;10 to 2;10, Levy discerned four distinct stages of acquisition. In Stage A, (1;10 to 2;0) the child used only the masculine plural

¹³ The rules of pluralization in Hebrew are complex, involving stress shift, vowel changes, and spirantization of stops as well as the addition of suffixes. The children Levy studied showed development in these areas as well.

ending *-im* to mark plurality, regardless of the gender of the noun. In Stage B (2;0-2;2) the child began using the feminine *-ot* plural morpheme on feminine nouns ending in *-a* but continued using the masculine *-im* on feminine nouns ending in *-t* and on all other nouns. In Stage C (2;2-2;5), the child began to expand *-ot* (sometimes also followed by an additional *-im*) to feminine nouns ending in *-at*, but continued to use *-im* for feminine nouns ending in *-et* and *-it*. Finally, in Stage D (2;5-2;10), the full complement of formal regularities was realized (albeit not always correctly) in the child's speech. All the evidence Levy presents points to the child using phonological cues to determine which plural morpheme would be used, and a gradual development of understanding as to which phonological cues required which plural form. Note that this pattern favors the masculine plural ending as the default plural, even to the point of adding *-im* to the end of the feminine plural in some cases. This offers additional evidence that Karmiloff-Smith's analysis of the masculine as a "default" gender may have cross-linguistic validity.

Levy found no evidence that her son made use of any semantic criteria either in the earlier stages (in Stage B for example, the word *íma* "mommy" became **ímaim* rather than the correct *imahót*) or in the later stages.

In the cross-sectional study, Levy tested 32 children age 2;2-2;10 using a Wug-type test (Berko, 1958) to elicit plural forms. Children were shown sixteen items, twelve of which were familiar objects, four of which were unfamiliar objects which were given nonce labels conforming to Hebrew phonology. Children were asked to name the object (or in the case of the unfamiliar object,

children were told the nonce name) and then shown a second, nearly identical object, and then asked what they saw before them. Since there were now two of the same object, the expected answer was a plural form.

As in the longitudinal study, children were significantly better at pluralizing masculine nouns and feminine *-a* nouns than they were at pluralizing feminine *-t* nouns. No significant semantic differences were found, either between nouns denoting animate and inanimate objects or between nouns denoting male and female objects (i.e., items which traditionally belong to either men or women)¹⁴. This suggests that once again, semantic gender information is less useful to the child than linguistic cues to gender, although this is not at all surprising given the very young age of the children in this study. Since children only begin to acquire an understanding of natural gender between the ages of 2 and 3 (see Chapter 2), it is entirely possible that some of the children in the experimental study (and certainly Levy's son during the time of the longitudinal study) had not yet acquired gender identity, and therefore it is moot to assess the importance of natural gender cues in their acquisition of the language.

One very interesting point Levy makes regards the persistence of certain types of errors. Levy notes that errors which are consistent with the general rules are more persistent than errors which contradict the general rules. For example, the feminine noun *beica* "egg" takes the irregular plural form *beicim* rather than

¹⁴ While items typically belonging to men and women do not themselves have natural gender, it is nonetheless not illogical that if children are using semantic criteria to assign gender, factors regarding who uses an inanimate object are likely to be taken account of as well. This seems to be the line of reasoning in using inanimate objects instead of pictures of actual males or females, although the author has not seen the issue of inanimate objects standing in for actual natural gender addressed in any of the research.

the expected **beicot*. Remember that the $-a > -ot$ rule is one of the most salient and therefore among the first of the pluralization rules to be learned. This salience makes the incorrect but regular **beicot* resistant to change. Deviant forms such as **simlim* (instead of *simlot*, from *simla* “dress”) which would be inconsistent with the regular paradigm, are eradicated first.

In comparing Levy’s and Karmiloff-Smith’s data it is important to note that the subjects in Levy’s cross-sectional study were a full year younger (2;2-2;10) than the youngest children in Karmiloff-Smith’s much larger study (3;2-3;11), and the boy in Levy’s longitudinal study was observed when he was even younger (beginning at age 1;10). Karmiloff-Smith’s youngest subjects had already acquired gender to a fair degree, in the sense that they were already past the point of making gender mistakes with familiar words. Thus Levy’s data may offer more insight into what children do in the earliest stages of acquiring gender. In addition, Levy’s study looked primarily at the children’s attempts to inflect real language whereas Karmiloff-Smith used only nonce words. Levy notes that the nonce words in her experiment were much more difficult for children to inflect than real words. Both sets of data however lead to very similar conclusions: children pay primary attention to the phonological end of the noun in determining gender. Levy goes further and states that semantic and syntactic distinctions are not utilized in early acquisition.¹⁵ It may be that children first familiarize themselves with the distribution of the phonological correlates of gender. Having

¹⁵Levy does not discuss syntactic cues in detail, but states several places that children appeared to be “insensitive” to syntactic cues (Levy (1983; 121)).

largely succeeded in this endeavor by the time they are about 3 (i.e., just after the age of Levy's subjects, just before the age of Karmiloff-Smith's subjects) they begin to turn their attention to syntactic correlates of gender (i.e, the distributional pattern of determiners and adjectives associated with the nouns), and finally to the semantic correlates of gender (i.e, the distributional pattern of masculine and feminine nouns denoting male and female entities). This would account for Levy's finding that syntax and semantics are irrelevant to early learners, and Karmiloff-Smith's finding that syntax is relevant but to a lesser degree than phonology.

3.3.3 *German*

Mills' (1986) study examines Levy's (1983) claim that natural gender is not salient to young children acquiring gender. She therefore looked only at the pronouns children use to refer to human beings where natural gender should be most salient. Specifically, she compared pronoun use by English- and German-speaking 3- and 4-year-olds. She found that German children were correctly inflecting pronouns (and determiners) for gender at a younger age than English-speaking children, suggesting that the pervasiveness of gender marking in German facilitated formation of nominal categories.

German has three genders, masculine, feminine and neuter. Gender marking is pervasive in German. Noun gender triggers agreement of associated determiners, adjectives, ordinal numbers and participles, as well as adjectival,

relative and question pronouns, and 3rd singular personal pronouns. English, of course, displays gender only in this latter category. In addition, all German nouns including inanimates are assigned gender. Most words for human and sex-differentiated animals take the gender of their natural sex, although this is not always the case (e.g., *mädchen* “girl”). All English gender is semantically based (that is, males are referred to with masculine pronouns, females with feminine pronouns.)

Mills tested English and German 3- and 4-year-olds by showing them six pictures of people engaged in various activities. Mills asked the children to give the person in the picture a name. Using this name to avoid cueing the children to the correct pronoun form, she asked the children questions about the pictures until the children had used two pronouns to refer to the person in the drawing. The following is an example of the type of conversation elicited from an English-speaking girl, age 4;8 (German conversations followed the same patterns):

(Experimenter shows a picture of a child wearing a dress who is in a field picking flowers.)

E: What's this? (pointing to the person) ¹⁶

C: It's a girl 'cause he's wearing a dress.

¹⁶ Although Mills (1986) attributes this exchange to a girl aged four years and eight months, other research indicates that pronoun are usually acquired much earlier by English-speaking children. Brener (1983) notes that correct production of English 3rd person pronouns precedes 1st and 2nd person pronouns. Furthermore, the children studied in Brener's comprehension study showed that children were even more attuned to gender distinction than they were to person distinctions (i.e., between speaker, addressee and other) choosing referents of the correct gender more often than referents of the correct person. However, this is an area where there may be considerable individual variation; Mills' subject may be a very late learner.

E: Can you think of a name for the girl?

C: Elizabeth.

E: What is Elizabeth doing in the picture?

C: She's pulling up flowers.

E: Where does Elizabeth put the flowers?

C: In the basket.

E: What is Elizabeth wearing?

C: a ... a hat, he got a hat on.

(Experimenter shows the next picture.)

Mills found that all German children consistently used correct gender-specific pronouns with over 90% accuracy. No differences could be observed between 3- and 4-year-olds, between boys' and girls' use of pronouns, or between pronouns used to refer to male and female referents. This indicates that German children had acquired the German pronominal gender system by age 3. English children however showed clear differences from the German-speaking children. English-speaking 3-year-olds were significantly below their German counterparts in their ability to consistently use the correct pronoun to refer to people. English-speaking boys seemed to have particular difficulty with female referents (61% accuracy), while girls made many mistakes with both sexes (42% accuracy for male referent, 55% for female referent). By age 4, the English-speaking children had almost caught up with German 4-year-olds, but boys still made many errors with female referents (70% accuracy). The fact that most errors were in the direction of using masculine pronouns for females rather than vice versa suggests that the English-speaking children overgeneralized the masculine pronoun.

Mills interprets these results as contradicting Levy's (1983) claim that young children do not notice the semantic distinction between males and females. However, this might be too strong a statement. First of all, the children in Levy's study had a mean age of 2;7 while the youngest children in Mills' study had a mean age of 3;8, an important difference. Furthermore, Mills' study specifically looks at children's use of personal pronouns in which natural gender determines the form, rather than pluralization of inanimate objects in which natural gender is non-existent and grammatical gender is determined primarily by morpho-phonological factors. Thus Mills' results do not directly contradict Levy's.

On the other hand, Mills' analysis of the reasons for differences between English and German children indicates that formal linguistic factors might be very much at play in German children's early acquisition of pronouns encoding natural gender distinctions. She notes that the pervasiveness of gender in German provides children with "more opportunity to see the regularities in the system and to learn to produce these forms accurately" (Mills, 1986; 41). English children, on the other hand, have virtually no similar opportunity. And, since articles (which are gender-inflected in German) are acquired earlier than pronouns, German children become aware of corresponding distributional patterns earlier than English children and might therefore transfer this early awareness to pronouns. Thus the lesson of Mills' study is that, even for forms where semantic distinctions are being encoded, the formal linguistic regularities seem to aid children in discerning the semantic distinctions, rather than the other way around.

Furthermore, Mills notes that the clear phonological distinction between masculine and feminine pronouns and articles in German probably aids in their acquisition. The German masculine (nominative) pronoun and article are *er* and *der* respectively, the feminine are *sie* and *die*. The phonological similarity in the final vowel is in stark contrast to the English. The English masculine nominative, accusative, and possessive pronouns are *he*, *him*, and *his*; the feminine are *she*, *her* and *hers*. The final vowel in the nominative is the same /i/ for both masculine and feminine, and a sibilant is characteristic of both the masculine possessive and the feminine nominative and possessive. Thus the English-speaking child has little phonological help in determining which forms are masculine or feminine. Given what Karmiloff-Smith (1979) and Levy (1983) have shown regarding the importance of phonology in French and Hebrew gender acquisition, the odds seem to be against English children learning the tiny gender system of their language easily.

3.3.4 Spanish

Perez-Pereira (1991b) demonstrated that Spanish-speaking children ages 4 to 11 rely more on formal linguistic properties to determine the gender of nouns than on semantic properties. This study is extremely interesting to consider due to the correlations with the study presented in Chapter 4, which also addresses the acquisition of gender in Spanish, although as L2 rather than L1.

Perez-Pereira's experiment is largely modeled on Karmiloff-Smith's (1979) experiments 7-11 (reviewed above); however, in a single experiment, he manipulated all three possible cues (semantic, phonological, and syntactic) that a child might use to determine gender. By using all possible combinations of these three cues, Perez-Pereira was able to ascertain the relative weight children gave to each type of cue, alone or in combination with other cues. The procedure was similar to that of Karmiloff-Smith (1979). He showed 160 children 22 drawings of imaginary beings, animals and objects, some of which showed natural gender, some of which did not. The pictures were all labeled with nonce names, some of which showed morpho-phonological cues to gender (i.e., ending in the vowels *-o* for masculine and *-a* for feminine) while some had gender-ambiguous word endings. Finally, the pictures and names were all introduced in a sentence which either gave a syntactic cue to gender (i.e., a masculine or feminine indefinite article) or used a non-agreeing cardinal number (*dos* "two"). These three cues were combined so that six of the items showed children only one type of cue to gender, six of the items showed two converging cues, and ten of the items showed two or three conflicting cues. Perez-Pereira then performed tasks similar to those performed by Karmiloff-Smith (1979) (such as covering up a picture) and asked questions about the pictures until the child used a gender-inflected adjective (generally a color) to describe the picture. The form of the agreement morpheme on the adjective showed which gender the child ascribed to that noun. By analyzing which cues and combinations of cues produced the most reliable

responses, Perez-Pereira was able to determine which type of cue Spanish-speaking children pay most attention to.

The first significant point to notice was the children's overgeneralization of the masculine, to a greater degree even than noted by Karmiloff-Smith. Masculine cues of all three types, presented singly or in convergence with another masculine cue, produced reliable masculine responses to the point of hitting a ceiling effect. Because of this, the forms which were predicted to elicit a feminine response provide the more interesting cases, and these produced much less reliable feminine responses. The cues which produce reliable feminine responses were the linguistic cues. The two items which showed either a feminine syntactic cue or a feminine morpho-phonological cue in isolation were marked as feminine by 65%-95% of the children, whereas the item which showed only feminine natural gender and no linguistic cues was reliably marked as feminine by 0%-20% of children. The difference in use between the linguistic and non-linguistic feminine cues was significant at $p < 0.001$. (The difference between the children's use of syntactic and morpho-phonological cues did not reach significance.)

The results show that children use formal linguistic cues more than semantic cues to assign grammatical gender. However, when semantic cues were presented converging with morphological or syntactic cues, children's reliability of gender assignment increased, suggesting that non-linguistic information was attended to, perhaps confirming and solidifying the linguistic information. This was especially evident when natural gender converged with the morpho-

phonological cue (95%-100%); convergence of the natural gender cue and the syntactic cue showed slightly less reliable result (only 80%-100%). Overall, there was no significant difference between syntactic and morpho-phonological cue use, but when two convergent cues were available, morpho-phonological cues appeared to be more salient than syntactic cues. When a linguistic and a non-linguistic cue were presented in conflict, the non-linguistic cue was largely ignored.

Interestingly, when there were contradictory cues, children tended to pay more attention to syntax than morpho-phonology. The 4- and 5-year-olds seemed to pay slightly more attention to morpho-phonological features, whereas older children seemed to pay slightly more attention to syntactic agreement, although overall there is little developmental difference between 4-year-olds' and 11-year-olds' cue use. Although the developmental differences here are slight, it is worth noting this might support Newport's (1988, 1990) theory that younger children can attend to only very small pieces of information (such as, perhaps, morpho-phonological endings), whereas older children can attend to larger chunks of information (such as, perhaps, agreement markers across word boundaries).¹⁷ We will return to this idea in Chapter 6.

Perez-Pereira interprets the results of this study similarly to Mills (1986) in suggesting that the regularity and productivity of an extensive grammatical gender system gives children an early opportunity to discover and practice the gender patterns. Therefore it should not be surprising that the formal linguistic

¹⁷ See Chapter 1 for further discussion.

factors are more salient and more useful to children. They have been attending to and discerning the formal linguistic factors longer than they have been able to clearly discern natural gender.

3.3.5 Icelandic

The evidence that grammatical gender is mapped first onto semantic gender is somewhat scantier. In English, where the gender system is a purely semantic one, children begin to use correct gender-specific pronouns only after they have passed the developmental milestone of recognizing and categorizing male and female (Mills, 1986). Mulford (1985) reports that Icelandic children more reliably select the correct referent of a gender marked pronoun when it refers to a person of that sex than to an object of that grammatical gender. However, Perez-Pereira (1991a) has criticized Mulford's study on the grounds that it included no sex-specific items that lacked grammatical gender. Since grammatical gender was always available as a cue in each situation, the children had no opportunity to show the relative weight they gave to the different cues, so their greater ability with natural gender-endowed objects shows rather a preference for two cues over only one.

3.4 Second language acquisition of gender

Research on the use of cues to gender in L2 acquisition is also available. This section will review several studies of gender acquisition among L2 learners. Some of these studies are replications of Karmiloff-Smith and Perez-Pereira's work outlined above, using L2 students as subjects rather than L1 speakers. In addition to shedding light on the use of linguistic and non-linguistic cues to gender, these studies also offer insight into how the instantiation of a grammatical gender feature in L1 influences the acquisition of gender.

Taylor-Brown (1984) attempted to replicate Karmiloff-Smith's (1979) French study, but with L2 immersion students rather than L1 speakers. Taylor-Brown studied English-speaking children in grades 3, 7 and 9 with various levels of French immersion education. She found that the English-speaking children did not seem to make use of phonological word endings in the same way that Karmiloff-Smith's Francophone subject had. In fact, the English speakers tended to ignore both the phonological cues and the syntactic cues in the form of indefinite articles and, instead, they tended to dramatically overgeneralize masculine determiners. Taylor-Brown repeated the experiments using the deictic determiners *ce* and *cette* with the thought that these might be more phonologically salient cues to gender than the indefinite articles. Although there was some statistically significant improvement in the English-speaking children's performance, their ability to correctly assign gender was still only 36%, with the strong tendency still to overgeneralize masculine gender.

Taylor-Brown also tested English-speaking children's use of natural gender as a cue to grammatical gender. In her replication of Karmiloff-Smith's Experiment 10 (showing obviously male and female beings paired with conflicting feminine and masculine phonological forms), she found that the children tended once again to ignore both potential cues to gender and respond using masculine gender. Only two children in the study consistently responded using a definite article which matched the natural gender of the picture, the rest used masculine articles 100% of the time. It would seem therefore that not only did the English-speaking children not attend to natural gender cues where the phonological cue was in conflict, in these cases they did not pay attention to either cue. Remember that Karmiloff-Smith also found a tendency for her older (10 and 11-year-old) Francophone subjects to overgeneralize the masculine definite article in this condition; at this age there was an increase in the use of the masculine article for male pictures with feminine noun endings, but not for female pictures with masculine noun endings, which were reliably marked as masculine by children of all age groups. However, the Francophone subjects did not overgeneralize the masculine definite article nearly to the extent that Taylor-Brown's Anglophone subjects did.

Carroll (1989) interpreted Taylor-Brown's research as evidence for a difference in how French gender is acquired by L1 and L2 learners. She argues that French L1 learners actually learn nouns and determiners together as "chunks" and then go through a period of reanalysis and split them apart, during which the correspondence between the noun endings and the determiner is realized. Carroll

suggests that L2 French learners learn the noun and determiner as discrete items to begin with, and that accounts for the difference in eventual attainment.

Taylor-Brown's results strongly suggest that speakers of English, a -gen language, do not use cues to gender in the same way that L1 speakers of French, a +gen language, do. However, as all of the subjects in Taylor-Brown's study were Anglophone, there is no comparison of how a +gen L1 might influence the acquisition of a +gen L2. The next study examines this question.

Bruhn de Garavito and White (2000) investigated the influence of L1 on L2 by looking at the acquisition of Spanish gender by two groups of French L1 students (ages 15 and 16, proficiency low and low-intermediate, respectively) who were studying Spanish at school. These subjects had only classroom exposure to the L2. Subjects were given sets of cards with pictures on them, from which they had to choose one picture to describe to the researcher. The researcher had to guess which picture the subject was describing. The naturalistic data obtained included many DPs, which of course required gender marking.

Bruhn de Garavito and White's results show that the low-intermediate group made significantly fewer errors in gender assignment than the low proficiency group. However, the types of errors made by the two groups are similar. For each group, accuracy was significantly higher when the DP contained definite articles than when it contained indefinite articles. Furthermore, both groups showed a significant tendency to overgeneralize masculine determiners to feminine contexts. That is, they mistakenly used masculine articles with feminine nouns more often than they used feminine articles with masculine nouns

(although there were a few individuals in the low-proficiency group who overgeneralized feminine to masculine contexts, especially with indefinite articles). Finally, when the noun was an entity with natural gender, both groups were less accurate in their gender agreement than when the noun was a sexless item (although this trend was significant only with the low-proficiency group).

Bruhn de Garavito and White then compare their own data on gender agreement accuracy with Hawkins (1998) results¹⁸. Hawkins assessed the accuracy of English L1 students learning French (that is, L1 is -gen but L2 is +gen). Hawkins's subjects had substantially more exposure to the L2 than Bruhn de Garavito and White's subjects; subjects were all college students, one group who had received their secondary education in a French immersion program, and the other who had averaged 10 years of classroom exposure and had at least 6 months immersion experience. The comparison showed that despite the difference in exposure and level of proficiency, both groups had similar levels of accuracy. Furthermore, subjects of both types of L1 tended to make more mistakes with indefinite than with definite articles, and to overgeneralize the masculine articles.

Bruhn de Garavito and White interpret these similarities as evidence that L2 gender acquisition difficulties are not attributable to the lack of a gender feature in L1, and that in fact the status of gender in L1 is irrelevant (Bruhn de Garavito and White, 2000). Franceschina (2003) on the other hand, points out

¹⁸ Hawkins does not include data on the presence or absence of a natural (sex) gender cue, so no comparison is made in this area.

that the data might have a different interpretation: the presence of the gender feature in their subjects' L1 may have led to an accelerated rate of acquisition of gender in the L2, allowing them to achieve a level of accuracy comparable to much more proficient learners whose L1 was –gen.

Each of the studies reviewed so far concentrate on data from groups where the L1 was either +gen or –gen, and comparison are only post hoc. Franceschina (2003) examined the extent to which the existence of a gender feature in L1 influences the acquisition of gender in L2¹⁹ by incorporating many more languages. She looked at the Spanish gender acquisition of adults who had all begun learning Spanish after puberty (all very advanced in their mastery of Spanish) and compared attainment by English L1 speakers (English being a –gen language) with attainment by L1 speakers of +gen languages (Arabic, Italian, Portuguese, French, Greek, and German). Unfortunately she did not have any–gen language other than English in her sample; nevertheless this study is one of the few to specifically address the issue by having English L1 and +gen L1 speakers complete identical tasks.

Franceschina's subjects performed a series of written and oral tasks, including cloze and multiple choice tests, grammaticality judgment tasks, novel word tasks (such as used by Karmiloff-Smith and Perez-Pereira) as well as others. She found that there was a very strong correlation between type of L1 (in terms of

¹⁹ Franceschina also looked at myriad other linguistic variables (such as case and number agreement), extra-linguistic variables (such as age of first exposure to Spanish and length of exposure), and task related variables (such as modality and data type). Since the majority of these are irrelevant to the study considered in this dissertation, not all of them will be reviewed.

+gen or -gen) and performance in all gender-related tasks. Speakers of a +gen L1 performed at the same level as native Spanish-speaking controls on all tasks, whereas English-speakers performed consistently below both native and +gen L1 speakers. Franceschina takes this as an indication that the ability to acquire gender is related to having that feature instantiated in one's L1. However, since English was the only -gen L1 for which she had subjects, this conclusion should remain tentative until similar results are obtained with additional -gen language groups. Nevertheless, Franceschina's study is extremely interesting and relevant to the study presented in this dissertation. The most relevant data for our purposes come from Tests 3 and 5, which were modeled on Perez-Pereira's (1991) study. The results of these two tests are considered below.

Test 3 was a cloze test that manipulated cues to grammatical gender using the same paradigm as Perez-Pereira. This test used real nouns instead of nonce words. Subjects had to choose a contextually appropriate noun from a list of distracters; for each item there were two contextually appropriate nouns, one of each gender. Each item in the test contained a combination of phonological, semantic, and syntactic cues to gender, after the manner of Perez-Pereira's test (Perez-Pereira did not include any items in which all three cues pointed to the same gender whereas Franceschina did; also, in this test Franceschina did not include items containing only semantic and phonological cues with no syntactic cue, although test 5 does include this relationship.)

As in all other results from this study, -gen (English) L1 speakers performed significantly worse than both native Spanish speakers and +gen L1

speakers, while natives and +gen L1 speakers did not differ significantly from each other. Not surprisingly, most of the mistakes in all three language groups were in the items where conflicting cues were presented, but it is striking that of the two conflicting conditions (phonology vs. syntax/natural and phonology vs. syntax only) by far the most mistakes were made in the first of these two conditions, where the natural sex of the referent was specified, and might have facilitated distraction from the spurious phonological cue. Instead, “it would seem that the semantic clues did not only fail to facilitate accuracy, but they made it more difficult” (Franceschina, 2003; 223).

Test 5 was essentially a replication of Perez-Pereira (1991), using the same words in the same conditions (see the previous section on L1 gender acquisition for a summary of the study and its findings) but this time with adults. She found that both syntactic and phonological cues are strong attractors for gender assignment, but syntactic cues are attended to more than either phonological cues or semantic cues in cases of conflict. Since both Perez-Pereira and Karmiloff-Smith indicated a slight developmental trend to attend more to syntax in their older subjects, it is not surprising to find this culminating in adults favoring syntactic over phonological cues. Feminine phonological forms were found to also be strong cues for gender assignment, suggesting possibly an adult awareness of masculine as an “unmarked” form and therefore less attractive as a cue for discerning gender. Semantic cues were relatively weak, and had most effect in combination with another converging cue.

In Test 5, the –gen (English) L1 speakers were significantly worse in assigning gender to nonce nouns than either their native Spanish- or +gen L1-speaking counterparts. Interestingly, however, the relative weight of the three cues did not vary by language group. All three groups showed a preference for using syntactic cues when they were available, a strong use of phonological cues when no syntactic cue was available, and only a weak use of the semantic (natural gender) cue. This suggests that despite a lower level of attainment in the area of gender, English speakers follow the same hierarchy of cues as native speakers. Perhaps therefore, it is not the weight these learners attribute to the cues which matters, but rather their ability to use this information to compute the appropriate gender form which lags behind native speakers’.

White et al. (2004) also include both –gen L1 (English) and +gen L1 (French) in their study of gender acquisition in L2²⁰ Spanish. All had had only classroom exposure to Spanish and none reported exposure prior to their mid teens. Subjects were grouped according to proficiency level into low, intermediate, and advanced cohorts. The study used two elicited production tasks designed to elicit naturalistic data, as well as a Spanish comprehension task including null nominals for which subjects had to identify the referents. There was also a vocabulary test. The results from all tests showed no distinction between the two L1 groups. Intermediate and advanced learners of both L1s showed results similar

²⁰ Many of the English speaking subjects were actually learning Spanish as their L3, L2 being French. White et. al address the issue of previous exposure to a +gen language in their data analysis and point out that there should be no difference between learning L2 or L3, as L1 remains constant. For purposes of this study the issue will not be addressed further.

to native Spanish speaking controls on gender accuracy. Low proficiency students of both groups made significantly more mistakes in gender accuracy, but their results did not differ significantly from each other – that is, both low proficiency groups, regardless of L1, showed difficulty in gender assignment. Again, White et al. interpret these results as indication that the presence or absence of grammatical gender in L1 is not a factor in L2 gender acquisition.

3.5 Current research questions

The majority of the research presented above examines gender acquisition in L2A by adults. Taylor-Brown's work, the only study addressing the gender acquisition of children, focuses on children with significant exposure to the L2 (French). The present study expands on the L2A research presented above by investigating acquisition of Spanish gender by children in a partial immersion setting after only a few months of exposure; specifically, this study investigates children's use of linguistic and non-linguistic (i.e. natural sex) cues to gender. It is hoped that by examining the children's early ability to make use of different types of cues, a better understand may be gained of how children begin to approach grammatical gender when gender is not manifested in their own L1.

The specific hypotheses to be tested in this study are the following:

1. Do English L1 children learning Spanish as L2 utilize natural gender to a greater degree than they use linguistic cues in determining the gender of nouns?

Null hypothesis: There is no difference in the use of linguistic and non-linguistic (natural sex) cues in the L2 acquisition of a +gen language by children of a –gen L1.

2. Do kindergarten and first grade children show any differences in the level to which they use linguistic and non-linguistic cues in acquiring gender?

Null hypothesis: kindergartners and first graders show no difference in the use of linguistic and non-linguistic cues.

3. Do boys and girls show any differences in the extent to which they use linguistic and non-linguistic cues in the acquisition of gender?

Null hypothesis: Boys and girls show no difference in the use of linguistic and non-linguistic cues.

4. Do L2 children use one gender more often as a default gender (in the way that L1 children are often shown to overgeneralize use of the masculine article to feminine nouns)?

Null hypothesis: There is no difference between the use of masculine and feminine gender by L2 children.

These hypotheses are addressed in the study presented here. Chapter 4 explains how this was done, and Chapter 5 addresses the results of the study in terms of each individual hypothesis.

3.6 Conclusion

To summarize, the research on L2 gender acquisition suggests that speakers of an –gen L1 can acquire gender to native-like levels. However, as most of this research has been conducted with older children and adults, the question is left open as to whether younger children may follow a path similar to L1 learners, or a path more like their adult L2 counterparts. Furthermore, with the exception of Taylor-Brown's (1984) study, the L2 research does not concentrate so much on attention to gender cues as it does on the influence of the L1. The study presented in this dissertation specifically addresses whether young children begin their acquisition of L2gender by attending to the same cues as children learning a +gen language natively. The next chapter will explain the study design and detail the experimental methodology.

Chapter 4

The purpose of the study was to investigate children's use of semantic and morpho-phonological cues in assigning noun gender. The study was therefore designed to be similar to those of Karmiloff-Smith (1979) and Perez-Pereira (1991), with a few modifications to address the specific hypotheses of interest, as well as to fit the subjects and their circumstances. This chapter describes the children who participated in the experiment, since an understanding of their linguistic abilities and their environment is important to understanding the design of the experiment. Next the experimental design and the type of data gathered from it are described in detail. Finally the method of analysis is explained.

4.1 Subjects

Initially 41 children participated in the study (although see below regarding subjects whose results were eliminated from analysis). The children were enrolled in the Kindergarten and First Grade Spanish Partial Immersion program at a recently opened International School in Seattle, Washington. This program provides approximately two and one-half hours per day of Spanish language immersion with native Spanish-speaking teachers from Spain and Mexico (kindergarten) and Bolivia (first grade). This study was conducted the first year that the International School was in operation; during this first year only

kindergarten and first grade had partial-immersion classes while the 2nd-5th grades continued with English only. For all the children in both classes, this was their first experience in an immersion program.²¹ At the time of the experiment, children had been in the program for approximately 7 months, for roughly 300 hours exposure to Spanish. Half of the children in each class had their Spanish sessions in the morning and their English sessions in the afternoon; the other half reversed the order. Science and math curricula were taught in Spanish; English was used for language arts and social studies curricula.

The children received no formal language teaching (that is, in which Spanish is the content as well as the medium of instruction). They had therefore had no formal instruction in the Spanish gender system, but had rather acquired their understanding of gender solely through natural exposure. Both teachers confirmed that neither the semantic nor the morpho-syntactic aspects of gender had been overtly pointed out to the children.

Both classes were observed twice, approximately three and two months before the experiment was conducted.²² Around this time the Kindergarten children were already using Spanish Det+N combinations with gender-appropriate articles in spontaneous speech (e.g., “Can I go to *el baño*” (*the bathroom*), and “Write it on *la pizarra*” (*the chalkboard*)), according to their teacher. This means

²¹ It later emerged that the first graders had received limited Spanish instruction during their kindergarten year, from the teacher who subsequently became their first grade teacher. This took place 3 times a week for about 30 minutes per session. The researchers were not aware of this at the time of the study.

²² The author had hoped to observe the children closer to the time of the experiment, however because of district-required assessments the experiment was postponed and the author was unable to observe the classes in the interim.

that the children appear to have a Spanish DP in their interlanguage grammar and possibly some concept of grammatical gender. During my observations of the class, children responded to the teacher's questions appropriately in Spanish when one to two word answers were required. Topics on which the children responded included greetings, counting and simple math tasks, colors, clothing, weather, dates and days of the week. Questions which involved longer answers (e.g., to the question “¿por qué...?”) evoked code-mixed responses (e.g., “*porque now we start over*”). First grade children were likewise observed employing code mixing (“this is *matematicas*, not *ciencias*”). Instances of spontaneous Spanish (rather than repetitions of the teacher’s speech) were few. Two children, asking to sing a song over again, spontaneously shouted “*¡Otra vez!*” and “*¡Uno más!*” Another child raised his hand to make the comment that “The song said *cantar* and we were *cantando!*”, which shows evidence that at least a few of the children were beginning to be aware of tense and aspect distinctions in Spanish. When children addressed me (e.g., to ask for a kleenex or for help tying their shoe) they did so in English. The inevitable muttering and whispering during circle time (not directed toward the teacher) and all free-play activity appeared to be entirely in English in both kindergarten and first grade.

These children were between the ages of 5 and 7 years²³, an age when gender identity and gender constancy are emerging or newly emerged, and thus natural gender is highly salient to them. For these reasons they were believed to

²³ Average age for kindergartners at the time of the study was 6.45 (slightly under 6 years 6 months). Average age for first graders is not available.

be at a perfect age and perfect stage of L2 acquisition for participation in this experiment.

4.2 Design

The experiment was conducted in a Training Phase and two Experimental Phases. The purpose of the Training Phase was to teach the children how to complete the experimental task. The four words used in the Training Phase and the 16 words used in Phase I were basic nouns for people and animals with which the children were expected to be reasonably familiar. The words used in Phase II were eight Spanish nonce words from Perez-Pereira (1991) used in a similar experiment with native Spanish-speaking children. The nonce words all represented invented animate (human- and animal-like) beings, which the children had not seen before.

The four words used in the Training Phase were a combination of animate and inanimate objects with ambiguous word endings (that is, not *-o* or *-a*). Phases I and II were designed to include all combinations of morpho-phonological and semantic cues to grammatical gender. The 16 words in Phase I were divided into four groups of four words, while the eight words in Phase II were divided into the same four groups but with only two words in each group. The cue combinations used in the four conditions are shown in Table 4.1.

Table 4.1: Combinations of cues to grammatical gender

Training Phase	Training	- natural gender, - morpho-phonological ending
Phase I and Phase II	Condition 1 (“natural only”)	+ natural gender, - morpho-phonological ending
	Condition 2 (“linguistic only”)	- natural gender, + morpho-phonological ending
	Condition 3 (“conflicting”)	+ natural gender, + morpho-phonological ending (conflicting)
	Condition 4 (“converging”)	+ natural gender, + morpho-phonological ending (converging)

Conditions 1 and 2 consist of words which offer *either* semantic (that is, natural sex) *or* morpho-phonological (that is, linguistic) cues to grammatical gender. By comparing the children’s categorization on these two conditions we hoped to see which type of cue children used more reliably in assigning gender. Condition 3 consists of words which offer *conflicting* cues to gender; that is, nouns that refer to naturally male entities but have typically feminine endings, and vice versa. This allowed the assessment of whether conflicting cues interfered with the child’s ability to correctly assign grammatical gender. Condition 1, 2 and 3 test the relative strength of the two types of cues. Condition 4 consists of words which offer *converging* semantic and morpho-phonological cues to grammatical gender; that is, nouns which refer to naturally male entities also had masculine endings, and nouns which refer naturally female entities had feminine endings.

This allowed the assessment of whether or not having two cues available enhanced the child's ability to correctly assign grammatical gender. Evidence of interference in children's ability to assign gender due to the conflicting cues in Condition 3, or enhancement of that ability due to convergent cues in Condition 4 in comparison with Conditions 1 and 2, would suggest that, even if children do pay more attention to one type of cue, they also are aware of and take note of the second type. No words or objects which lack both semantic and morpho-phonological cues were included in either experimental phase; since this experiment was designed to test the relative weight children give the two different types of cues, such objects would not be enlightening.

For the words in Condition 2, pictures of animals were chosen rather than pictures of inanimate objects. Levy (1983a) found that children ages 2;2-2;10 performed better inflecting nouns denoting animate than inanimate objects when the words were nonce nouns. Although Levy did not find the same difference between animates and inanimates in common nouns, it was preferable to avoid introducing the possible confound that children might treat nouns denoting animate and inanimate objects differently. Also, when collapsing the nonce and real data, no confound is introduced by collapsing both animate and inanimate object. It might be argued that by using animals instead of inanimate objects, the additional confound was introduced of the children's potential understanding that animals also have natural gender. This is a valid concern; however, none of the pictures of the animals showed any sexual attributes, and it was believed, based

on experience, that children within this age range would be unlikely to think of animals in the same terms of “he” and “she” that they associate with people.

The words in Condition 3 are morpho-phonologically invariable but can take either grammatical gender, the choice of gender depending on the sex of the referent. For example *el piloto* refers to a male pilot while *la piloto* refers to a female pilot; *el artista* refers to a male artist whereas *la artista* refers to a female artist. This means that, unlike in Conditions 1, 2, and 4, whichever gender a child ascribes to these words, the chosen article with the noun will form a grammatical DP. Words were carefully selected to be familiar to the children and at the same time avoid too blatantly contradicting sex-role stereotypes the children may or may not have, causing gender confusion of an unrelated sort.

The words used in the Training Phase and Phase I are listed in Table 4.2. The nonce words used in Phase II are listed in Table 4.3. Note that there was no additional Training Phase preceding Phase II. In addition, no “Grammatical Gender” is defined for the nonce words, since they do not actually have a formal gender assignment.

Table 4.2: Words and gender cues in Phase I

Condition	Natural Gender	Suffix	Word 'gloss'	Gramma-tical Gender
Training	∅	∅	<i>bebe</i> 'baby'	M
	∅	∅	<i>elefante</i> 'elephant'	M
	∅	∅	<i>flor</i> 'flower'	F
	∅	∅	<i>leche</i> 'milk'	F
Condition 1 ("natural only")	M	∅	<i>hombre</i> 'man'	M
	M	∅	<i>rey</i> 'king'	M
	F	∅	<i>mujer</i> 'woman'	F
	F	∅	<i>cantante</i> 'singer'	F
Condition 2 ("linguistic only")	∅	M	<i>caballo</i> 'horse'	M
	∅	M	<i>perro</i> 'dog'	M
	∅	F	<i>vaca</i> 'cow'	F
	∅	F	<i>cebra</i> 'zebra'	F
Condition 3 ("conflicting")	M	F	<i>artista</i> 'artist'	M
	M	F	<i>debntista</i> 'dentist'	M
	F	M	<i>piloto</i> 'pilot'	F
	F	M	<i>modelo</i> 'model'	F
Condition 4 ("converging")	M	M	<i>niño</i> 'boy'	M
	M	M	<i>abuelo</i> 'grandfather'	M
	F	F	<i>niña</i> 'girl'	F
	F	F	<i>enfermera</i> 'nurse'	F

Table 4.3: Words and gender cues in Phase II

Condition	Natural Gender	Suffix	Word
Condition 1 ("natural only")	M	∅	<i>tanten</i>
	F	∅	<i>pifar</i>
Condition 2 ("linguistic only")	∅	M	<i>carepo</i>
	∅	F	<i>anira</i>
Condition 3 ("conflicting")	M	F	<i>tiba</i>
	F	M	<i>lacato</i>
Condition 4 ("converging")	M	M	<i>milipo</i>
	F	F	<i>capola</i>

The purpose of Phase II was to avoid the possibility that children would have memorized the Det+N combinations used in Phase I, showing complete mastery of the forms and offering no insight into which cues were more useful to them. Since the children had not seen the drawings or heard the words in Phase II, it was hoped that this would allow us to see what children would do with new words when first confronted with them. However, Levy (1983a) notes that children's ability to inflect nonce words often lags behind their ability to inflect real language. In her study, children were significantly better at pluralizing animate nonce nouns than inanimate, although the same was not true for familiar nouns. She also found that very young children (ages 2;0-2;2) offered different types of answers when asked to inflect real and nonce words. Children almost always attempted to inflect real words even when they did not know the correct form, but did not try to inflect the nonce words; rather, they either repeated the uninflected word or remained silent (Levy, 1983a). Levy suggests that "a task

that involves nonce words is handled by the children as different in kind from an equivalent one with familiar words” (Levy, 1983a; 119). In fact, it might be argued that when children first meet any new word, the initial process used to determine inflection is different than that used when the word becomes familiar (since all new words are, in effect, nonce words to a child). In the present study, it was hoped that by employing nonce words, we would be able to see that early process more clearly than in the real-word task. It was anticipated that the children in this study, being several years older than those in Levy’s, would attempt to inflect the nonce nouns using whatever system they had intuited for Spanish gender. To avoid the dichotomy Levy (1983a) noticed between animate and inanimate nonce nouns, all the nonce nouns in this study were animate.

The nonce words were eight words randomly chosen from Perez-Pereira’s twenty-two nonce words, so as to be maximally sure that legitimate Spanish phoneme combinations were used. However the words were not used in the same configurations; that is, Perez-Pereira used *tanten* as word with neither natural nor phonological cue, whereas in the present study *tanten* was used to name a drawing showing a natural gender cue.

It might be objected that a syntactic cue should have been included in the experiment. For example, it would have been possible to follow a paradigm akin to Perez-Pereira’s (1991) or Karmiloff-Smith’s (1979) in which either an indefinite article showing agreement or a number (e.g., *dos*) which does not show agreement was systematically presented along with the pictures. There are several reasons this was not done. First of all, in none of the literature has it been

shown that children use syntactic agreement as a primary cue to gender. Some of Perez-Pereira's older subjects appeared to give primacy to syntactic cues over phonological, but these results were not consistent and in some cases not statistically significant. Perez-Pereira interpreted his results as showing the primacy of phonology as a cue, and the primacy of linguistic cues over non-linguistic. Karmiloff-Smith (1979) and Levy (1983), too, interpreted their results as indicating that phonological cues outweighed other linguistic cues in determining gender. Thus it was decided that for this experiment the most promising avenue of inquiry would include the morpho-phonological and semantic cues rather than syntactic. In addition, the children in the immersion program where the research was conducted had only a few months exposure to Spanish. They were at a very low level of proficiency, a level at which they produced only some simple vocabulary and a few formulaic expressions. It seemed that a larger experiment testing all three variables would be too complex for the children and might yield uninterpretable results. For these reasons the present experiment tested only two variables.

4.3 Materials

20 drawings of the objects, animals and people listed in Table 4.2 were used in the Training Phase and Phase I, and 8 drawings of the entities listed in Table 4.3 were used in Phase II. The drawings in the Training Phase and Phase I were created by the author by enlarging clip-art pictures from the Internet, tracing

them, and coloring them. The drawings in Phase II of the experiment were drawn by me in a similar fashion to those shown in Perez-Pereira (1991a & b) and Karmiloff-Smith (1979). Drawings for Conditions 1, 3 and 4 depicted whimsical beings with secondary sexual characteristics to indicate natural gender. Drawings for Condition 2 showed animal-like beings with no secondary sexual characteristics (samples of drawing are in Appendix A). Each of the drawings was colored, mounted on black construction paper, and laminated.

4.4 Procedure

The experiment was conducted over the course of a week and involved 41 children, 23 kindergartners (10 boys, 13 girls) and 18 first graders (eight boys, nine girls). Children came to the experiment session from their Spanish classroom, except for the last two Kindergarten children who for scheduling reasons had to come from their English classroom. Children were seated at the end of a low table. A native Spanish-speaking researcher (named Lola) and a native English-speaking researcher (the author) sat on either side of the child. Each researcher spoke to the child only in her native language. After explaining the general purpose of the session to the child in English (“to learn more about how children learn a new language”), the author showed the child each of the four drawings from the Training Phase and Phase I individually and asked him or her to say in English what they thought that drawing was. Next Lola asked the child

to tell her the Spanish word for each drawing if they knew it, and if they didn't, not to worry because she would tell them what it was.

Doing separate English and Spanish labeling of the drawings proved to be a necessary adjustment from the original design of the experiment, which had called for the children to begin by labeling each drawing in Spanish if they could, in English if they could not. The first few children seemed to have trouble determining the level of detail being requested in drawings of people, which were the majority of the items. For example, the drawing designed to depict a "grandfather" or *abuelo* could be labeled "person," "man," "old man," or "grandfather"; the drawing designed to depict a "man" or *hombre* could be labeled "person," "man," "dad," or some occupation ("lawyer" was one child's contribution). The cognitive load of determining the level of detail they were meant to attend to in each picture as well as find the appropriate word in their Spanish vocabulary seemed too high, and a few children seemed to give up by about halfway through the labeling task and just waited for us to tell them what the drawing was. Having the children name the drawings first in English (on the pretext that "sometimes my drawings aren't very clear") seemed to help them with the Spanish labeling.

After the children had viewed the pictures and heard them labeled twice (once in English, once in Spanish with Lola), the English-speaking researcher told the children in English that Lola would now introduce two friends, Elo and Lala. Elo and Lala were a lion and a tiger puppet. After the children had had a chance to greet the puppets and pet them, the children were told (in English) that Elo and

Lala spoke only Spanish but sometimes they disagreed on how things should be said in Spanish, especially whether they should use *el* or *la* when talking about the things in the pictures. The English-speaking researcher asked the child to look at each picture and listen to how Elo and Lala spoke about it, and then to decide which animal was speaking about it right in Spanish, and give the picture to that puppet. The research held the pictures and handed them to the child while the two puppets labeled the pictures in Spanish, this time using the definite article. Elo always used the article *el* and Lala always used the definite article *la*.

The Training Phase pictures were the first four pictures used in each labeling task and in the Phase I task. Lola had the puppets “repeat” the Spanish label with both articles. Thus for the picture of a “boy,” the child heard the puppets (through Lola) say “¡La niño!” “¡El niño!” If at any point the child was unsure which puppet to give a picture to, the English-speaking researcher suggested that he or she put it to one side and decide later.

Upon completing Phase I, the child was praised for helping the puppets. The English-speaking researcher then introduced the Phase II drawings by saying that these were pictures of people who came from other planets, and since they aren’t from this planet they can’t really be called “people.” The children were told (in English) that they would hear what each kind of “person” was called in Spanish, but that Elo and Lala would probably disagree again about whether they should use *el* or *la* to talk about them. They should give the picture to the puppet that was talking about it right.

Throughout the experiment, the lion puppet Elo consistently used *el* while the tiger puppet Lala consistently used *la*. This was done to avoid confusing the child (as well as the experimenter manipulating the puppets!)

The reason for asking children to categorize the drawings, rather than using a design which requires spontaneous production of the articles and nouns, was to avoid confounding the experiment with performance factors. Children in the early stages of immersion programs rarely use the target language spontaneously (Blanco-Iglesias, Broner and Tarone, 1996). Although the teacher reported that the children did use some Spanish vocabulary in the class (and this had been observed as well), the change from classroom to experimental setting might cause some children to become shy and produce less Spanish than they ordinarily do, or than they are capable of. In addition, grammar mistakes are common in immersion children's spontaneous speech. Therefore it would not be surprising for children to make mistakes in gender despite knowing the correct form. It might likewise be true that even when children do not *know* the correct form, they have an idea (or, in Carroll's (1989) terms, a "rule of thumb") of how to figure out the correct form, but forget to employ it. Since the interest of this study is not in children's production but in what they know about the gender system of Spanish, this design filtered out performance factors as much as possible and allowed the researchers to see only the children's conscious perception of grammatical gender.

All the children said they found the tasks enjoyable, and most truly did seem to enjoy the session. However, we found that children tested toward the end

of the week were less engaged, seemed more tired, and participated less fully.

Several children seemed disturbed by external factors: one girl told the researchers that her friend wouldn't play with her on the playground at recess, and one child was clearly diverted by some commotion near the testing area. The majority of the children completed the task, only one did not. Five of the children completed the experiment by pointing always to one puppet or the other. In four of those cases the child pointed to the same puppet for both Phase I and Phase II, in one case the child switched puppets at phase II. After the testing, three first graders' (one boy and two girls) and four kindergartners' (two boys and two girls) scores were eliminated from the analysis for the following reasons:

- BF18 – commented as soon as he heard the task: “Oh I know the lion will be right all the time,” then gave all pictures to the lion (masculine)
- GF39 – pointed only to lion (masculine)
- GF29 – informed us that her parents speak a lot of Spanish with her (also evidenced in the fact that she got everything right).
- BK31 – spoke Thai at home, pointing only to tiger (feminine)
- BK30 – mom and aunt at home spoke Spanish
- GK13 – studying Chinese at weekend Chinese school, pointed only to tiger (feminine) in phase I, only to lion (masc) in phase II.
- GK27 – did not look at pictures, pointed only to lion (masculine).

This reduced the total number of subjects for analysis to 34. This number included 20 kindergartners and 14 first graders; 15 were boys and 19 were girls.

4.5 Analysis

In the analysis of the results, the real words were initially analyzed separately from the nonce words. However, the decision was made to collapse the

two sets of words for a number of reasons. First, the means for each condition for all groups were remarkably similar. Secondly, with only two words per condition (Phase II), statistical consultants advised that statistical tests were unlikely to reveal true difference in the means for the nonce words. It was suggested that by collapsing the data into a single set of six words, the data would better facilitate the statistical procedure being used.

To determine whether the presence of the different gender cues increased the children's likelihood of correctly identifying the gender of the noun, it was necessary to discern whether the mean number of times the children answered correctly in one condition was significantly different from the mean number of times they answered correctly in the other conditions. Therefore a proportion was calculated of correct answers each child gave (that is, the number of times each child gave a picture to the correct puppet) for each of the four conditions. If the child gave all six pictures in any condition to the correct puppet, the child received a score of 1.0 for that condition. If the child gave three of the six pictures to the correct puppet, the child received a score of 0.5 (for two correct pictures 0.33; for four correct picture 0.67, etc). A sample of this calculation for one child is in Table 4.4 below. This child (a boy, first-grader, ID#3) gave 4 of the 'natural only' pictures to the correct puppet, none of the linguistic pictures to the correct puppet, 3 of the "conflicting cues" pictures to the correct puppet, and 3 of the "concord cues" pictures to the correct puppet. Thus, he received a score of 0.667 for the natural condition, 0 for the linguistic group, 0.5 for the conflicting

the converging groups. The table of calculations (converted to percentages for clarity) for all 34 children is provided in Appendix B.

Table 4.4: Example of calculations per condition for a single child

	Natural	Linguistic	Conflicting	Converging	Sex	Year	ID
Right	4	0	3	3	B	1	3
Wrong	2	6	3	3			
Percent	67%	0%	50%	50%			

The proportions calculated for all children were then analyzed using a series of paired t-tests to assess whether children's responses in the four conditions differed significantly. A paired t-test was used because this test accounts for the correlation between any given child's answers. That is, the response that "Child A" gave for the word *hombre* is more closely correlated to the response of the same child for the word *enfermera* than it is to "Child B"'s response for *hombre*. A paired t-test between any two given conditions is the same as calculating the difference between those two conditions for each child and then performing a one-sample t-test to find out whether or not this difference is equal to zero.

Tests were initially conducted on the full set of data (minus the children whose results were eliminated). To be conservative, a two-tailed rather than a one-tailed p value was taken. After examining the results of these tests,

subsequent t-tests were run on the data stratified by age (all kindergartners, all first graders) and by sex (all boys, all girls). All t-tests were done using Microsoft Excel.

4.6 Conclusion

This chapter has explained the details of how the experiments was designed and conducted. It has also explained the method of analysis. In Chapter 5, the results of the experiment are presented and analyzed, and in Chapter 6 some implications of the results are discussed.

Chapter 5

5.1 Introduction

This chapter presents the results from the study. First, restated below are the hypotheses tested. These will be expressed both as a “question” to be answered, and then as a null hypothesis which will be supported or rejected.

1. Do English L1 children learning Spanish as L2 utilize natural gender to a greater degree than they use linguistic cues in determining the gender of nouns?

Null hypothesis: There is no difference in the use of linguistic and non-linguistic (natural sex) cues in the L2 acquisition of a +gen language by children of a –gen L1.

2. Do kindergarten and first grade children show any differences in the level to which they use linguistic and non-linguistic cues in acquiring gender?

Null hypothesis: kindergartners and first graders show no difference in the use of linguistic and non-linguistic cues.

3. Do boys and girls show any differences in the extent to which they use linguistic and non-linguistic cues in the acquisition of gender?

Null hypothesis: Boys and girls show no difference in the use of linguistic and non-linguistic cues.

4. Do L2 children use one gender more often as a default gender (in the way that L1 children are often shown to overgeneralize use of the masculine article to feminine nouns)?

Null hypothesis: There is no difference between the use of masculine and feminine gender by L2 children.

Each question will be examined in turn. All the results presented in this chapter include the data from both Phase I (real words) and Phase II (nonce words).

5.2 Use of linguistic vs. natural cues to gender

The first question is whether English L1 learners of Spanish L2 utilize natural or linguistic cues to grammatical gender, or both. To answer this question, it is instructive first to examine the means correct (the mean number of times children handed a picture to the correct puppet) for all four conditions.

Table 5.1 shows the mean correct for all children, as well as stratified by age group and by sex. For each condition, if the children were using the available cue or cues to gender, they were likely to get the gender right (that is, give the picture to the correct puppet) whereas if they were not attending to the cue, they were likely to be guessing, and have only a 50-50 chance of getting the answer

right. Therefore, in the conditions where children were utilizing the available cues, the mean correct for those conditions should be higher than the means in the conditions where they were not using the available cues.

Table 5.1: Means correct for all groups and conditions.

	Natural	Linguistic	Conflicting	Converging
All Children	48.5%	54.9%	<i>45.6%</i>	63.2%
Kindergartners	50.0%	58.3%	<i>42.5%</i>	68.3%
First Graders	<i>46.4%</i>	50.0%	50.0%	56.0%
Boys	<i>41.1%</i>	52.2%	<i>41.1%</i>	63.3%
Girls	54.4%	57.0%	<i>49.1%</i>	63.2%

For each group of subjects, the condition in which they were most successful is in bold; the condition(s) in which they were least successful is in italics.

Table 5.1 shows that all groups of children were most successful in assigning gender when presented with two converging cues. They were least successful in the conflicting condition (with the exception of first graders who were least successful in the Natural condition; and boys, whose lack of success in the Conflicting was matched by their lack of success in the Natural condition). When both natural and linguistic cues pointed to the same gender assignment, children had the greatest chance of correctly identifying the gender of the picture, but when the two cues indicated different genders, the children are least likely to correctly identify the gender of the picture. This suggests that the children in this study had begun to be aware of gender agreement rules and had at least a rudimentary understanding that both natural and linguistic cues may be useful in assigning gender.

With respect to the single-cue conditions, for all groups without exception, children were more successful assigning gender in the linguistic condition than in the natural. This suggests that, contrary to Hypothesis 1, these children are not transferring their limited gender category from English L1 into Spanish L2. Rather, they appear to be paying more attention to the linguistic cue than to the natural cue to gender, similar to what is known for L1.

Table 5.1 presents the means for all subjects in all four conditions, both as a full cohort and stratified by age and sex. The next section examines more closely the data for the full group.

5.3 Statistical analysis of complete data

The findings indicate that children used the morpho-phonological cues more than natural cues, yielding a higher mean of correct gender answers in the linguistic condition than in the natural condition. In order to determine whether this difference was statistically significant, the data from the linguistic condition were compared with the natural condition. In addition, the data from the “conflicting cues” and “converging cues” conditions were also compared. Remember that in these two conditions, the children had access to both natural and linguistic cues for each picture. If children attended to both cues at least somewhat, we should expect to find that children are significantly more successful at discerning the correct gender when both cues point to the same gender than when the two cues indicate different genders. Table 5.2 shows these

statistical comparisons for all children. (The complete set of statistical tables for all comparisons can be found in Appendix C.)

Table 5.2 Differential use of cues (all children)

Condition	t-stat	p value (two-tailed)
Natural only vs Linguistic Only	-1.14084	0.262
Conflicting vs Converging	-3.35304	0.002

Table 5.2 shows that the difference in the children's use of linguistic or natural cues presented in isolation does not reach statistical significance. This means that neither cue shows evidence of the children using it to a degree greater than predicted by chance. However, there is a significant difference between how accurately children were able to answer in the case of conflicting vs converging cues ($p < 0.05$). Children had a much higher correct average in the case of converging cues (grand mean: 63.2% correct) than in the case of conflicting cues (45.6% correct). This finding is very important. It may be that a single cue of either type gives children at this stage of acquisition only a vague idea of what grammatical gender the noun is, so they are less able to make the choice. It may be that they are aware of the validity of both cues, and at this level of learning they really need the two converging cues to be able to make a gender determination. This would explain why the difference between the two isolated cues fails to reach significance – even if the children tend more toward the grammatical cue, they are less sure of the correctness of doing so, but when both cues are present they are more secure in their judgment.

On the other hand, it cannot really be ascertained from these tests whether the significant difference between conflicting and converging conditions is because the two converging cues helped the children, or because the conflicting cues hindered the children in correctly identifying gender. In order to test this, the results in the natural condition were compared with the conflicting condition. Remember that in the case of the conflicting cues, the linguistic cue points to the “wrong” gender. If the children are attending more to the linguistic cue than to the natural cue, then we would expect them to mis-assign gender in this case, because the linguistic cue they are using is “throwing them off”. In both natural and conflicting conditions the natural cue is the correct indicator of gender. If the children perform significantly worse in the conflicting condition than in the natural condition, it may be that the linguistic cue is grabbing the children’s attention at the expense of the natural cue. On the other hand, if the presence of two cues over one is helpful to the children, we should expect to find significant differences between the converging cues on one hand and *either* cue presented in isolation on the other hand. These tests are presented in Table 5.3.

Table 5.3: Differential use of cues (all children)

Condition	t-stat	p value (two-tailed)
Natural only vs Conflicting	-1.1408	0.568
Natural only vs Converging	-2.8770	0.007
Linguistic only vs Converging	-1.5257	0.137

Table 5.3 shows no significant difference between the “natural only vs. conflicting, so we cannot make any determination as to whether the conflicting cues were a hindrance or the converging cues were a help. However, there is a significant difference between the natural vs. converging cue conditions. This suggests that when the natural cue was available, the added presence of the grammatical cue pointing the children in the same direction significantly helped the children make the correct gender assignment. However, the difference between the linguistic only condition and the converging condition is not significant, which suggests that when the morpho-phonological ending is present, the addition of the natural cue does not significantly help the child assign gender. Both of these findings suggest that the children are paying more attention to the linguistic than to the non-linguistic cue.

Thus, although the evidence is not conclusive, we have reason to believe that we should reject the null hypothesis expressed in 1 above. It seems that the children in this study are paying more attention to linguistic cues than to natural cues in determining gender. Although the differences are not great, the statistical difference between the natural only and the converging cues suggests that the linguistic cue may be more helpful to children than the natural cue.

It is possible that analyzing the data for all children, rather than for specific groups of children, is masking significant effects for one group by lumping it together with the non-significant effects found in the other group. To address this possibility, let us now turn to the question of age-related differences, analyzing the kindergartners and first graders separately.

5.4 Do kindergartners and first graders make different use of cues?

As discussed in Chapter 1, much research into L2A indicates that “younger is better” in learning a second language. In this study we are presented with a unique opportunity to investigate this issue. All of the children in this study had received naturalistic exposure to Spanish for the same amount of time, because both kindergartners and first graders were beginning the year for the first time in an immersion program. Therefore the only difference between the two groups (assuming input of equal quality from their respective teachers) is their age at first exposure to the L2 (that is, the age at which they began the program)²⁴. The means for the two age groups are presented again in Table 5.4 below for convenience.

Table 5.4: Kindergartners and first graders mean accuracy in each condition.

	Natural	Linguistic	Conflicting	Converging
Kindergartners	50.0%	58.3%	42.5%	68.3%
First Graders	46.4%	50.0%	50.0%	56.0%

We now examine whether the kindergartners and first graders used the cues in similar ways. As pointed out above, both groups achieved the highest level of accuracy in the converging cues condition. The kindergartners follow the same trend seen above with the whole group – they were somewhat better at

²⁴ Subsequent to completing the study and the writing of this dissertation, the author was made aware that the first graders had in fact had limited exposure to Spanish during their kindergarten year. After inquiries, it was decided that the limited and non-naturalistic nature of the exposure did not materially affect the results of this study.

assigning gender given linguistic cues in isolation than natural cues in isolation, and they did the worst on the conflicting condition.

What is interesting is that the first graders do not appear to follow this trend at all. While they did do best on the converging condition, note that their mean correct in this condition is the lowest of all groups (refer to Table 5.1). In addition, the first graders worst scores are found in the natural-only condition rather than on the conflicting condition. In both the linguistic condition and the conflicting condition they were exactly at 50%, which suggests they may not have been paying much attention to any of the cues.

Again however, the observational data is submitted to statistical evaluation to determine whether these groups show reliably more use of the natural or linguistic cue. Table 5.5 shows the kindergartners' scores on the comparisons explained above.

Table 5.5: Kindergartners' differential use of cues; p values shown in bold are significant ($p < 0.05$).

Kindergartners: Condition	t-stat	p value (two-tailed)
Natural only vs Linguistic Only	-1.26892	0.220
Conflicting vs Converging	-3.58713	0.002
Natural only vs Conflicting	1.072085	0.297
Natural only vs Converging	-2.7729	0.012
Linguistic only vs Converging	-1.2554	0.225

The results in Table 5.5 show that the kindergartners are using the linguistic and non-linguistic cues to gender to different extents. Although kindergartners appear to attend more to linguistic cues than to natural cues, as

evidenced by their higher mean correct in the linguistic condition, the difference between the two conditions fails to reach significance. However, kindergartners' difference in accuracy between the conflicting and converging conditions is extremely significant. The kindergartners appear to be aware of the validity of both natural and linguistic cues. Presented with one cue in isolation they may be less able to make a distinction than when presented with two cues, and in the face of two conflicting cues they may be unsure which cue to attend to. Alternatively, they may have been attending to one cue which was aiding them in the converging condition and/or hindering them in the conflicting condition. Given the highly significant results in comparing the conflicting to converging conditions, the latter of these may be more likely.

In order to test this, kindergartners' use of the natural cue was compared with their success in the conflicting condition, where they should attend to the natural cue and ignore the linguistic one (the same comparison that was made for the group as a whole). If they were attending to the linguistic cue when it was present, we would expect that they would be significantly more successful in the natural condition, where there was no linguistic cue to interfere than in the conflicting condition.

Table 5.5 shows no significant difference between the linguistic vs. converging conditions, so we cannot make any determination as to whether the conflicting cues were a hindrance or the converging cues were a help. On the other hand, there is a significant difference between the means correct in the converging condition compared with the natural-only condition. It appears that

when the natural cue was available, the added presence of the grammatical cue pointing the children in the same direction significantly helped the children make the correct gender assignment. However, the difference between the linguistic only condition and the converging condition is not significant, which suggests that when the morpho-phonological ending is present, the addition of the natural cue does not significantly help the child assign gender. Both of these suggest that the children are paying more attention to the linguistic than to the non-linguistic cue, although the difference between these two cues in isolation does not show significance.

Examining the results for the first graders, a very different picture emerges. None of the comparisons of the different cue conditions show any significance at all for the first graders. The results of the same test using the first graders correct for each condition are shown in Table 5.6.

Table 5.6: First graders differential use of cues

First grade: Condition	t-stat	p value (two-tailed)
Natural only vs Linguistic Only	-0.35488	0.728
Conflicting vs Converging	-0.89142	0.389
Natural only vs Conflicting	-0.4932	0.630
Natural only vs Converging	-1.169	0.263
Linguistic only vs Converging	-0.83602	0.418

Given these results, it appears that the kindergartners are attending to gender cues in a way that first graders are not. Based on these results, first graders

cannot be shown to be attending to any cues at all. Their gender assignments can only be regarded as random.

These results are even more interesting when placed within the context of discussion of age-related factors in L2A, discussed in Chapter 1. Some previous research has suggested that younger children acquire an L2 at a faster rate than older children and adults. The findings here indicate that, for gender, this may be because the younger age children approach the acquisition task through linguistic cues, whereas older children, even only a year older as these are, have begun to approach the task in a different way (although what strategy the older children might be using is impossible to say from these data). The results here do not explicitly support any of the potential theories expounded either in favor or against a critical period, but they do show that in gender acquisition, as in other areas of L2A, younger children appear to have the advantage over older children. Further research involving many more children however, would be required to assert more than this with certainty.

5.5 Do Boys and Girls make different use of cues?

While it is not a central tenet of this dissertation that there is a difference between boys and girls in how they learn language, it is not impossible that this is the case. Because the topic of this study related to the use of natural gender cues, and because (as mentioned in Chapter 2) the children involved in this study are of an age at which gender permanence is becoming established, it seemed interesting to take advantage of the fact that both boys and girls took part in the test. The two

groups' means are repeated below from Table 5.1 (here as Table 5.7) for convenience.

Table 5.7: Mean accuracy per condition for boys and girls.

	Natural	Linguistic	Conflicting	Converging
Boys	41.1%	52.2%	41.1%	63.3%
Girls	54.4%	57.0%	49.1%	63.2%

In Table 5.7 we observe that both boys and girls were more accurate in assigning gender in the linguistic condition than in the natural condition, suggesting that both groups attend more to the phonological cue than to the sex of the referent. Furthermore, both boys and girls were most accurate in their gender assignments in the converging condition, and were least accurate in the conflicting condition. Interestingly however, boys' low accuracy in the conflicting condition was matched in the natural only condition. They are the only group in this study to have performed as inaccurately in the natural condition as in the conflicting condition.

Table 5.8: Boys' differential use of cues

Boys: Condition	t-stat	p value (two-tailed)
Natural only vs Linguistic Only	-1.17734	0.259
Conflicting vs Converging	-3.34664	0.005
Natural only vs Conflicting	-1.2	1
Natural only vs Converging	-2.75092	0.016
Linguistic only vs Converging	-1.23359	0.238

Table 5.9: Girls' differential use of cues

Girls: Condition	t-stat	p value (two-tailed)
Natural only vs Linguistic Only	-0.38838	0.702
Conflicting vs Converging	-1.78022	0.092
Natural only vs Conflicting	0.753277	0.461
Natural only vs Converging	-1.36646	0.189
Linguistic only vs Converging	-0.8914	0.384

Table 5.8 and Table 5.9 show that boys follow the pattern seen previously for kindergartners and for the group as a whole. Their greater accuracy in the linguistic condition than in the natural condition, while observable in the mean, does not reach significance. However they perform significantly better at assigning gender when presented with two converging cues than when presented with two conflicting cues ($p < 0.005$). Once again, there is no significant difference between the linguistic vs. converging conditions, so again we cannot make any determination as to whether the conflicting cues were a hindrance or the converging cues were a help. However, we again find a significant difference between the natural vs. converging cue conditions. Once again, this suggests that when the natural cue was available, the added presence of the grammatical cue pointing the children in the same direction significantly helped the children make the correct gender assignment. However, as with the kindergartners and the group as a whole, the difference between the linguistic condition and the converging condition is not significant, indicating that when the morpho-phonological ending is present, the added natural cue does not significantly assist boys in assigning

gender. Again, we interpret both of these facts to mean that boys are paying more attention to the linguistic than to the non-linguistic cue.

While boys show significant differences in several areas, Table 5.9 shows that girls do not. In none of the comparison tests do the girls' scores show any significant difference between use of the different types of cues in any combination. Although their means trend in the same direction as other groups' and in the same direction as the group as a whole, it cannot be asserted based on the data here that girls are paying more attention to one gender cue over the other, nor that they are better able to assign gender in the face of two converging cue than in other conditions. At this point, the girls' gender assignment, like the first graders' appears to be random.

5.6 Do children overgeneralize one gender?

Researchers investigating L1 acquisition of gender all stress that young children begin by over-extending one gender into all context. As discussed in Chapter 3, Karmiloff-Smith (1979), Levy (1983a & b), and Perez-Pereira (1991) all showed a tendency for their subjects to extend masculine determiners to non-masculine nouns. Taylor-Brown (1984) and Bruhn de Garavito and White (2000) both indicated similar findings, especially with indefinite articles. It was therefore a question whether these L2 children would follow a similar pattern of extending the masculine to feminine contexts more than vice versa.²⁵

²⁵ Remember that, of the seven children whose scores were eliminated from the analysis, five of them pointed only to one puppet or the other throughout the test. Two pointed always to the tiger

To test this, percentage correct for masculine nouns and the percentage correct for feminine nouns were compared, for all words and for the words in each condition. Means correct for masculine nouns and feminine nouns, by condition, are shown in Table 5.10.

Table 5.10: Means correct for masculine nouns vs feminine nouns.

	<i>masc</i>	<i>fem</i>
Overall	54.8%	50.2%
Natural	50.0%	47.1%
Linguistic	54.9%	52.0%
Conflicting	42.2%	49.0%
Converging	72.1%	53.0%

The data in Table 5.10 shows that the mean correct for masculine nouns was higher than the mean correct for feminine nouns, in all conditions but the conflicting condition. However, in the linguistic condition and the converging condition the children were still over 50% correct with feminine nouns. This suggests that the children were not overgeneralizing masculine definite articles to feminine contexts. Indeed, when these means are submitted to statistical analysis, the only condition in which the difference in gender accuracy between the masculine and feminine is significant is the converging condition. Table 5.11 incorporates Table 5.10 and shows t-stats and p values.

puppet (feminine) and three pointed always to the lion puppet (masculine). The author believes this was due to discomfort with the testing situation more than due to using a default gender for all nouns. Four of the five were eliminated for additional reasons as well (see Ch. 4). None of the seven were included in the analysis of default gender.

Table 5.11: Statistical difference for masculine and feminine nouns, by condition

	<i>masc</i>	<i>fem</i>	t Stat	p value (two tailed)
Overall	0.548	0.502	1.390407	0.167
Natural	0.5	0.471	0.414938	0.681
Linguistic	0.549	0.520	0.50148	0.619
Conflicting	0.422	0.490	-1.2276	0.228
Converging	0.721	0.530	2.770973	0.009

Based on the statistics above, there is no indication that the children were significantly overgeneralizing masculine responses to feminine contexts. The significant difference in accuracy between masculine and feminine in the converging condition is the only indication that the children might have been treating one gender of nouns differently from the other. Since again, children were 53% accurate using feminine definite articles with feminine nouns, this cannot be due to overgeneralization of the masculine.

Nonetheless, the extremely high accuracy (in comparison with other conditions) in assigning masculine in the converging condition deserves note. The opinion of the author is that the high percentage of correct gender assignment in this case is being driven by the nonce word *milipo*, referring to a being with antennae wearing something like a tuxedo. This word evoked fully 85% correct gender assignment, by far the most accurate responses of all the pictures in the entire study. Why this should have been so is a mystery. It is unlikely to be related to the natural cue in this picture, as it is the author's opinion that the drawing looked no more masculine than the pictures of, say, the man *hombre* or the

grandfather *abuelo*; besides, it was noted above that the natural cue appears to have been less attended to by most groups than the morpho-phonological ending. The author has no explanation for the high accuracy of responses for *milipo*; at this point nothing more can be asserted than that some combination of cues, phonology, and possibly the drawing conspired to make this drawing extremely salient for the masculine determiner.²⁶

5.7 Conclusion

Chapter 5 has presented the data from the study. The trends in the data have been examined and submitted to statistical tests. The results show that the L2 Spanish learners in this study paid more attention to the morpho-phonological (that is, linguistic) cue to gender than to the natural sex of the reference (a non-linguistic cue). While this trend is evident across the groups, the only differences in cue use which are statistically significant are those seen in the cohorts of boys and kindergartners. Also, there is no significant difference between use of linguistic and non-linguistic cues in isolation. Cue combinations appear to be more salient to the children. Chapter 6 will discuss these findings and their implications in light of the research presented in Chapter 1, Chapter 2, and Chapter 3.

²⁶ There is one interesting side note. In Perez-Pereira's data from native Spanish speakers, there are three words which elicited responses of 100% accuracy from all age groups, from four-year-olds all the way up to eleven-year-olds: *milipo*, *carepo*, and *nepo*. The author is not aware of any theory of phonology which asserts that final morphemes with a voiceless bilabial onset hold any predictive power toward gender, and would not presume to advance such an explanation here, but the phenomenon is somewhat striking.

Chapter 6

6.1 Single cues vs. multiple cues

The results of this study show that on the whole, the English L1 children who participated in this study did tend to pay attention to both the morpho-phonological cue and to the natural, non-linguistic cue. This is most clearly shown by the significant results obtained in comparing the conflicting and converging cue conditions. This is the comparison in which we should expect to see the greatest differences between the means if the children were attended to both cues. This comparison showed a significant difference for conflicting and converging conditions for the group as a whole, as well as for kindergartner and for boys; however, not for first graders or girls.

The results also suggest that as a group, the children in this study used the linguistic cue to a greater degree than they used the natural cue. This trend is clearest again among kindergartners and boys, although unfortunately, the results of the comparison between the two single-cue conditions do not reach significance for any of the cohorts examined. The more convincing data comes from the comparison between the natural condition and the converging condition, where boys and first graders were significantly more successful at assigning gender when the linguistic cue was added to the natural only cue (as were all the children as a group, although to only a near-significant level). Given the very interesting results from this study, it is disappointing that they do not show a clearer trend in one direction or the other. Many of the comparisons where we

would have expected to find significance (especially, the comparison between natural and linguistic conditions) show no significance. Why, since the mean levels of accuracy in the different conditions appear to show a clear trend, are they not more significant? The following discussion offers a few thoughts on this.

The lack of significant difference between the single-cue conditions may indicate the children have not yet reliably acquired gender to the extent that they are comfortable making overt metalinguistic decisions about it. The presence of a single cue may not yet be a strong enough indicator of gender for these children. Comparing Perez-Pereira's results, he points out that "the greater the number of converging clues that exist, the more easily children seem to determine the gender of nouns" (Perez-Pereira, 1991; 584). If this is true for Spanish-speakers with native comfort in the language, it may be even truer for these non-native speakers. Note also that this may have been the first time the children were ever asked to focus on the form of the language they were learning; focus on form is explicitly not a goal of the immersion program. It may be the case that for non-native speakers, 6 months partial-immersion exposure to Spanish is too little to make them comfortable focusing overtly on form in the single cue conditions.

6.2 Age- and sex-related differences

The fact that the trend toward using the linguistic cue over the natural cue was more pronounced in kindergartners than in first graders suggests that age does indeed have an effect on the acquisition of the feature +gen. Kindergartners appear to be attending to the linguistic cue in a way that first graders, only a year

older, are not. While it impossible in this case to firmly attribute this difference to age (there were other differences between the groups, notably they had different teachers and the absolute equivalence of the input for each group cannot be vouched for), it is nonetheless extremely suggestive that younger children are acquiring gender in a manner similar to L1 learners of a +gen language, whereas the first graders are not. It is similarly interesting that the boys, but not the girls, showed a significant trend toward attending to the linguistic cue over the non-linguistic cue. Is it possible that boys and kindergartners share an ability or a characteristic that girls and first graders do not?

As mentioned in Chapter 1, Newport (1988, 1990) has suggested that a lower developmental level of cognitive ability may actually allow very young children to attend to smaller bits of language (that is, the morpho-phonological endings) in a way that isn't facilitated by a higher level of cognitive development. If attending to smaller chunks of language is beneficial for acquisition, it should be no surprise that the very youngest of us, the least cognitively developed, should be the best at learning language. What's interesting is that there is long-known evidence that boys do develop cognitively much later than girls do (Fenson et al., 1994; Eliot, 2003) and that this difference is almost certainly genetic (Huttenlocher et al., 1991). It may be that we are seeing the effects of early cognitive development on children's linguistic ability. The girls in this study, who likely initially mature at a faster rate, may indeed be losing the ability to focus only on the small pieces of language that give the boys an edge in the

acquisition process, the same edge enjoyed by the group of kindergartners (both boys and girls).

Examining such a hypothesis would require comparing the difference between boys and girls at each grade level (that is, comparing the percentage of correct responses by kindergarten boys with those of kindergarten girls, and of first grade boys with those of first grade girls.) The hypothesis would suggest that kindergarten boys would show a greater attention to linguistic cues than kindergarten girls, that first grade boys would show greater attention to linguistic cues than first grade girls, and that kindergarten girls and first grade boys might pay similar attention to the different cue conditions. Given the small number of subjects participating in this study, this was not possible at this point.

Note however, that while boys' use of the cues appears to pattern in a way that suggests attention to linguistic cues, it is actually the girls who are more accurate in their gender assignments. Their level of accuracy in the single cue conditions actually approximated the kindergartners, and their level of accuracy in the natural only condition is the highest of all the cohorts (see Table 5.1, above). This suggests that while the girls' attention to the natural and linguistic cues may not be differentiated to the same degree as boys' and kindergartners', girls nonetheless are beginning to acquire the gender system to a higher level than the boys. It may be that some groups develop other strategies which compensate for their focus on the important linguistic cues, such as Carroll's (1989) "rules of thumb."

6.3 Familiarity with test stimulus

Another possibility for the lack of significance in the results is that the words that formed the basis of the test were too unfamiliar to the children. The words in Phase I were chosen because they were representative of the groups of words we hoped to test, and also because we hoped the children would be familiar with them. Most children were very familiar with the four training words (although most thought the word for 'flower' was *flore* or *flora*, rather than *flor*). No records were kept as to which article children chose for the training words as the training words did not clearly exhibit natural or morpho-phonological gender. Unfortunately, children were much less familiar with the Phase I experimental words than had been hoped. First graders tended to be familiar with the animal words in the linguistic only condition, except *cebra*, as well the words in the converging condition except *enfermera* (and most children initially said the nurse was a doctor, suggesting that the old stereotype of nurses being women and doctors being men is largely irrelevant to the youngest generation). Kindergartners tended to know only the words *niño*, *niña*, and occasionally *perro*. This was disappointing as it was hoped that words such as *hombre*, *mujer*, and *enfermera* would be familiar to the children. The words *cantante*, *rey*, and the words in the conflicting condition were assumed to be unknown. However it was also known that most of the children were familiar with the verb *cantar*, and hopefully the cognate nature of the nouns in the conflicting condition made them at least somewhat memorable for the 15 minutes that it was necessary for the

children to understand them. Only one child knew the word 'model' even in English.

6.4 Generic use of masculine forms in Spanish

Another question is whether the generic use of masculine forms in Spanish to refer to both masculine and feminine beings may have been as salient to the children as the two genders themselves. One boy (BF18) who knew 5 words chose the article *el* for all 24 words in the experimental phase²⁷ (although he did utter *la vaca*, *la niña*, and *la enfermera* (*enfermera* was not one of the words he knew)). Another boy (BF15) clearly knew the words *niño* and *niña* but informed us that either article was acceptable for *niña*. This claim leads one to wonder if generic use of masculine forms, which is common in Spanish, might have led children astray not only with the word *niña* but with the concept of gender as a whole. During the author's first observation of the Kindergarten class the teacher was trying to get the children's attention by saying "*Niños, niños*". One child must have questioned her because she was then heard to say (in Spanish) "When I say *niños* I mean both *niños* and *niñas*." Thus it is possible that children may still be confused about when masculine signifies a neutral gender and when it is gender specific (although note that there was no statistical significance found in the children's approach to masculine vs. feminine nouns; see also Roca (2000) for an interesting treatment of the Spanish masculine as generic). Therefore if

²⁷ BF18's results were not included in the statistical analysis.

children's answers were influenced by the generic use of masculine, the effect on these tasks was very subtle.

6.5 Avenues for future research

The fact that the statistical results were only moderately strong in this study indicates that a good next step would be to continue to assess children's gender acquisition with subsequent groups of children after the same amount of exposure as the children in this study had experienced (possibly working more closely with classroom teachers to determine better prompts for them to respond to). The benefit would be to provide further tokens of each gender assignment and amass a larger data pool, hopefully leading to a more robust set of data on which to perform the statistics. Again, splitting the subject group into cohorts of kindergartners and first graders, as well as boys and girls, would allow corroboration of the results from this study and potentially allow researchers to extend the results to see whether there are indeed statistical differences between gender assignments in conditions of single gender cues. This would provide evidence for one of two possibilities. If the addition of further data creates a more robust data set which in turn reveals significant differences between the linguistic only and non-linguistic only cue conditions, this could be taken as an indication that the size of the data set was too small in this study and it masked the importance of the individual cues. If there is still no difference between the means in the single cue conditions, the conclusion could more confidently be drawn that a single cue to gender does not provide enough evidence to children in order to

make a gender determination, and that it is the convergence of multiple cues which aids children in gender acquisition.

To that end, it would also be interesting to include the syntactic cues to gender, as was done in the studies reviewed in Chapter 3. Syntactic cues were not used in this study because the determination was made that children with only a few months exposure to Spanish probably had not acquired the language to the level that would be necessary to use that syntactic cue, given the assertions of previous research on gender acquisition. However, if, as we suspect, children are better able to make use of two cues to gender than they are of one cue, it may be that the syntactic cue might be used in conjunction with other available cues, even though on its own it may not provide enough evidence for gender.

The children in this study were in the very earliest stages of L2 acquisition. Thus the results obtained in this study can only be indications of the way children begin to approach the language acquisition task. The results are important because they show the children's ability to assign gender even before they begin to produce much Spanish. An important next step would be to return to test these children's production of gender as soon as they become comfortable enough with the language to utter Spanish phrases in limited elicited contexts. A further possibility would be to use a test such as Bruhn de Garavito and White's (2000) freer elicited production task, in which more conversational production surfaced a larger number of nouns and accompanying determiners (both definite and indefinite) and adjectives.

In this way, by following a group of children over a period of time and adjusting the test method to their current level of acquisition, it would allow us to create a more complete picture of the early acquisition of gender, the pattern of gender development, and, given enough opportunity to follow the children to fluency, the final state of gender in their L2.

A question which has not yet been considered is whether attention to linguistic cues over natural cues actually aids the L2 student in acquiring gender. Reexamining the mean correct by cohort and by condition (Table 5.1 is repeated below as Table 6.1 for convenience) it is clear that the kindergartners not only appear to use the linguistic cue more than the natural cue (although remember that this difference was not significant), they also have the highest level of accuracy in the linguistic condition and the converging condition, the two conditions in which attention to the linguistic cue would be most helpful. First graders, by contrast, achieved only 50% accuracy in the linguistic condition and only 56% in the converging condition, the lowest level of any group. On the surface, the fact that the age cohort which paid more attention to the linguistic cues also achieved greater gender accuracy overall would seem to suggest that attention to linguistic cues aids L2 learners in the same way that it aids L1 learners.

Table 6.1: Means for all groups and conditions.

	Natural	Linguistic	Conflicting	Converging
All Children	48.5%	54.9%	45.6%	63.2%
Kindergartners	50.0%	58.3%	42.5%	68.3%
First Graders	46.4%	50.0%	50.0%	56.0%
Boys	41.1%	52.2%	41.1%	63.3%
Girls	54.4%	57.0%	49.1%	63.2%

On the other hand, it was pointed out in Chapter 5 that, despite the lack of significance in girls' use of the two types of cues, girls approach kindergartners' accuracy level in the linguistic cue condition, and are nearly identical to boys' accuracy in the converging condition. They also achieved the highest level of accuracy in the natural only condition. It is entirely possible that the girls have developed some other means of assessing gender which does not rely so much on attention to linguistic cues. If this is the case, the next step should be to discover what that strategy might be. Again, this would require a longer, more in-depth study which looks both at children's initial learning patterns and their later patterns in terms of attention to gender cues, and tracks developing proficiency as attention to cues develops.

References

- Berko, J. 1958. The child's learning of English morphology. *Word* 14, 150-177.
- Bialystok, E. 1997. The structure of age: in search of barriers to second language acquisition. *Second Language Research* 13, 116-137.
- Bialystok, E. and K. Hakuta. 1999. Confounded Age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second Language Acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Lawrence Erlbaum Associates. 161-181.
- Birdsong, D. and M. Molis. 2001. On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language* 44. 235-249.
- Blanco-Iglesias, S., J. Broner, and E. Tarone. 1995. Observations of language use in Spanish immersion classroom interactions. In L. Eubank, L. Selinker, M. Sharwood Smith (Eds.), *The current state of interlanguage: Studies in honor of William E. Rutherford*. Philadelphia: J. Benjamins. 241-254.
- Braine, M. D. S. 1987. What is learned in acquiring word classes – a step toward an acquisition theory. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brener, R. 1983. Learning the deictic meaning of third person pronouns. *Journal of Psycholinguistic Research* 12, 235-262.
- Bruhn de Garavito, J., and L. White. 2002. The second language acquisition of Spanish DPs: The status of grammatical features. In J. M. Liceras and A. T. Perez-Leroux (Eds.), *The acquisition of Spanish morphosyntax: the L1/L2 connection*. Dordrecht: Kluwer Academic Publishers.
- Butt, J. and C. Benjamin. 1988. *A New Reference Grammar of Modern Spanish*. London: Edward Arnold.
- Cacciari, C., M. Carreiras and C. B. Cionini. 1997. When words have two genders: anaphor resolution for Italian functionally ambiguous words. *Journal of Memory and Language* 37, 517-532.
- Carreiras, M., A. Garnham and J. Oakhill. 1993. The user of superficial and meaning-based representations in interpreting pronouns: evidence from Spanish. *European Journal of Cognitive Psychology* 5, 93-116.

- Carroll, S. E. 1989. Second Language Acquisition and the computational paradigm. *Language Learning* 39, 535-594.
- Carroll, S. E. 1999. Input and SLA: Adults' sensitivity to different sorts of cues to French gender. *Language Learning* 49, 37-92.
- Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Clarke, M. A., A. Losoff, and M. D. McCracken, J. Still. 1981. Gender perception in Arabic and English. *Language Learning* 31, 159-167.
- Cochran, B. P., J. L. McDonald, and S. J. Parault. 1999. Too smart for their own good: the disadvantage of a superior processing capacity for adult language learners. *Journal of Memory and Language* 41, 30-58.
- Corbett, G. 1991. *Gender*. Cambridge: CUP.
- Corbett, G. and N. Fraser. 2000. Gender assignment: a typology and a model. In G. Senft (Ed.), *Systems of Nominal Classification*. Cambridge: CUP.
- Curtiss, S. 1977. *Genie: A psycholinguistic study of a modern day "wild child"*. New York: Academic Press.
- Flege, J. E. 1999. Age of learning and second language speech. In D. Birdsong (Ed.), *Second Language Acquisition and the Critical Period Hypothesis*. Mahwah, NJ: Lawrence Erlbaum Associates. 101-131.
- Eliot, L. 2003. *What's going on in there? How the brain and mind develop in the first five years of life*. New York: Bantam.
- Fenson, L., P. Dale, J. Reznick, E. Bates, D. Thal, and S. Pethick. 1994. *Variability in early communicative development*. Monographs of the Society for Research in Child Development, Vol. 59, No. 5, Serial # 242.
- Franceschina, F. 2003. The nature of grammatical representations in mature L2 grammars: the case of Spanish grammatical gender. Unpublished PhD dissertation, U. of Essex.
- Gass, S. and J. Ard. 1984. Second language acquisition and the ontogeny of language universals, in W. Rutherford (ed.), *Language Universals and Second Language Acquisition*. Amsterdam: John Benjamins. 33-68.

- Guiora, A. Z. B. and W. R. Acton. 1979. Personality and language behavior: a restatement. *Language Learning* 29, 193-204
- Guiora, A. Z., B. Beit-Hallahmi, R. Fried., and C. Yoder. 1983. Language Environment and Gender identity attainment. *Language Learning* 32, 289-304.
- Harley, B. 1984. Age as a factor in the acquisition of French as a second Language in an immersion setting. In R. W. Andersen (Ed.), *Second Languages: A Cross-Linguistic Perspective*. Rowley, MA: Newbury House. 143-163.
- Harris, J. W. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22, 27-62.
- Hawkins, R. 1998. The inaccessibility of formal features of functional categories in second language acquisition. Paper presented at the Pacific Second Language Research Forum, Tokyo. March 1998.
- Hawkins, R, and C. Y.-H. Chan. 1997. The partial availability of universal grammar in second language acquisition: the 'failed functional features hypothesis.' *Second Language Research* 13, 187-226.
- Hernandez-Chavez, E. 1977. The acquisition of grammatical structures by a Mexican-American child learning English. Unpublished PhD dissertation, UC Berkeley.
- Herschensohn, J. 2000. *The second time around; Minimalism and L2 acquisition*. Amsterdam: John Benjamins Publishing Co.
- Huang, J. 1970. A Chinese child's acquisition of English syntax. Unpublished MA Thesis, UCLA.
- Huttenlocher, J., W. Haight, A. Bryk, M. Seltzer, and T. Lyons. 1991. Early vocabulary growth: Relation to language input and gender. *Developmental Psychology* 27, 236-48.
- Johnson, J. and E. Newport. 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 39, 215-258.
- Karmiloff-Smith, A. 1979. *A functional approach to child language*. Cambridge: CUP.
- Keller-Cohen, D. 1979. Sytematicity and variation in the non-native child's acquisition of conversational skills. *Language Learning* 29, 27-44.

- Kellerman, E. 1978. Giving learners a break: native language intuitions as a source of predictions about transferability. *Working Papers in Bilingualism* 15, 59-92.
- Kellerman, E. 1995. Age before beauty; Johnson and Newport revisited. In L. Eubank, L. Selinker, M. Sharwood Smith (Eds.), *The current state of interlanguage: Studies in honor of William E. Rutherford*. Philadelphia: J. Benjamins. 219-231.
- Krashen, S. 1973. Lateralization, language learning and the critical period: Some new evidence. *Language Learning* 23, 63-74.
- Krashen, S. 1982. Accounting for child-adult differences in second language rate and attainment. In S. Krashen, R. Scarcella, and M. Long (Eds.), *Child-adult differences in second language acquisition*. Rowley, MA: Newbury House. 202-226.
- Lado, R. 1957. *Language across cultures*. Ann Arbor: University of Michigan Press.
- Larsen-Freeman, D., and M. H. Long. 1991. *An introduction to second language acquisition research*. New York: Longman Press.
- Lenneberg, E. H. 1967. *Biological foundations of Language*. New York: John Wiley.
- Levy, Y. 1983a. The acquisition of Hebrew plurals: the case of the missing gender category. *Journal of Child Language* 10, 107-121.
- Levy, Y. 1983b. It's frogs all the way down. *Cognition* 15, 75-93.
- Lucy, J. 2000. Systems of nominal classification: a concluding discussion. In G. Senft (Ed.), *Systems of nominal classification*. Cambridge: CUP.
- MacWhinney, B. 1978. *The acquisition of Morphology*. Monographs of the Society for Research in Child Development, No. 174).
- Maratsos, M. P. 1988. The acquisition of formal word classes. In Y. Levy, I. M. Schlesinger, and M. D. S. Braine, *Categories and Processes in Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates. 31-44.
- Maratsos, M. P. and M. A. Chalkley. 1980. The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In K. E. Nelson (Ed.), *Children's Language*, vol. 2. New York: Gardner Press, Inc.

- Martinez, I. M. and M. Shatz. 1996. Linguistic influence on categorization in preschool children: a cross-linguistic study. *Journal of Child Language* 23, 529-545.
- Mills, A. E. 1986. Acquisition of the natural gender rule in English and German. *Linguistics* 24, 31-45
- Mulford, R. 1985. Comprehension of Icelandic pronoun gender: Semantic versus formal factors. *Journal of Child Language* 12, 443-453.
- Mullen, M. K. 1990. Children's classification of nature and artifact pictures into female and male categories. *Sex Roles* 23, 577-587.
- Newman, A., D. Bavelier, D. Corina, P. Jezzard, and H. Neville. 2002. A critical period for right hemisphere recruitment in American Sign Language processing. *Nature Neuroscience* 5, 76-80.
- Newport, E. 1988. Constraints on learning and their role in language acquisition: Studies of the acquisition of American Sign Language. *Language Sciences* 10, 147-172.
- Newport, E. 1990. Maturation constraints on language learning. *Cognitive Science* 14, 11-28.
- Oliphant, K. 1997. Acquisition of grammatical gender in Italian as a foreign language (Net Work #7) HTML document. Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
<http://www.lll.hawaii.edu/nflrc/NetWorks/NW7/>
- Perez-Pereira, M. 1991a. Semantic versus formal theories of gender acquisition: A criticism of Mulford's study. *Archives de Psychologie* 59, 3-16.
- Perez-Pereira, M. 1991b. The acquisition of gender: what Spanish children tell us. *Journal of Child Language* 18, 571-590.
- Roca, I. M. 2000. On the meaning of gender. *Hispanic Research Journal* 1, 113-128.
- Schachter, J. 1974. An error in error analysis. *Language Learning* 24, 205-14.
- Schlesinger, I. M. 1971. Production of utterances and language acquisition. In D. Slobin (Ed.), *The Ontogenesis of Grammar*. New York: Academic Press.

- Schlesinger, I. M. 1988. The origin of relational categories. In Y. Levy, I. M. Schlesinger, and M. D. S. Braine, *Categories and Processes in Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates. 121-178.
- Sera, M. D., C. A. H. Berge, and J. C. Pintado. 1994. Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development* 9, 261-292.
- Sharwood-Smith, M. and E. Kellerman. 1986. Cross-linguistic influence in second language acquisition: An introduction, In Kellerman E. and M. Sharwood-Smith (Eds.) *Cross-linguistic influence in second language acquisition*. Oxford: Pergamon.
- Slavoff, G. R. and J. S. Johnson. 1995. The effects of age on the rate of learning a second language. *Studies in Second Language Acquisition* 16, 1-16.
- Slobin, D. I. 1966. Developmental psycholinguistics. In F. Smith and G. A. Miller (Eds.), *The Genesis of Language*. Cambridge, Mass.: MIT Press.
- Slobin, D. I. 1985. Cross-linguistic evidence for the language-making capacity. In D. I. Slobin (Ed.), *The Cross-linguistic Study of Language Acquisition Vol. 2: Theoretical Issues*. Hillsdale, NJ: Lawrence Erlbaum Associates. 1157-1256.
- Snow, C. and M. Hoefnagel-Hohle. 1978. The critical period for language acquisition: Evidence from second language learning. *Child Development* 49, 1114-1128.
- Taylor-Brown, K. 1984. The acquisition of grammatical gender by children in French immersion programs. Unpub. Masters Thesis, University of Calgary, Alberta.
- Teschner, R. V. and W. M. Russell. 1984. The gender patterns of Spanish nouns: An inverse dictionary-based analysis. *Hispanic Linguistics* 1, 115-132.
- Valian, V. 1986. Syntactic categories in the speech of young children. *Developmental Psychology* 22, 562-579.
- Weinreich, U. 1953. *Language in Contact*. The Hague: Mouton.
- Werker, J. and R. Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7, 49-63.

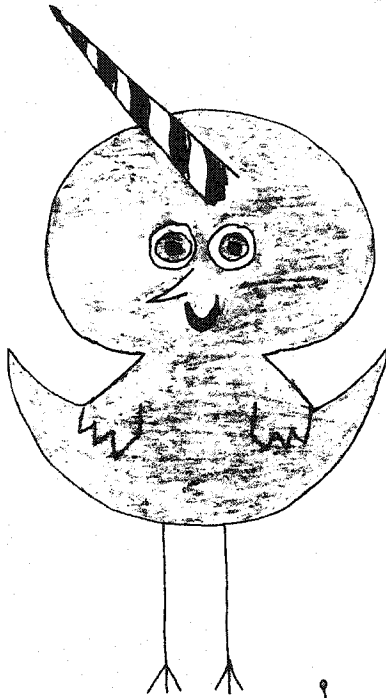
White, L., E. Valenzuela, M. Kozłowska-Macgregor, and I. Leung. 2004. Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics* 25: 105-113.

Zobl, H. 1980. Developmental and transfer errors: their common bases and (possibly) differential effects on subsequent learning. *TESOL Quarterly* 14, 469-79.

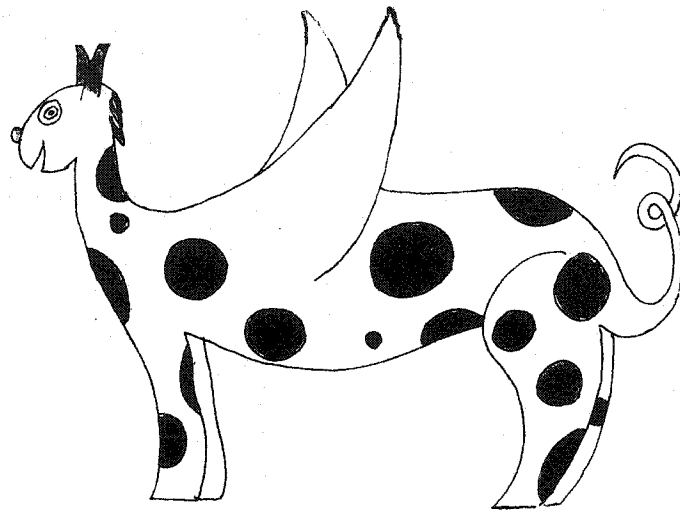
Appendix A

Samples of pictures used in the experiment.

Nonce Words



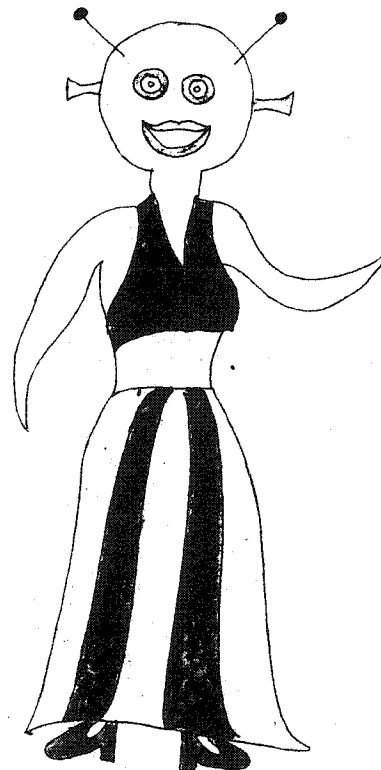
(la) anira



(el) carepo

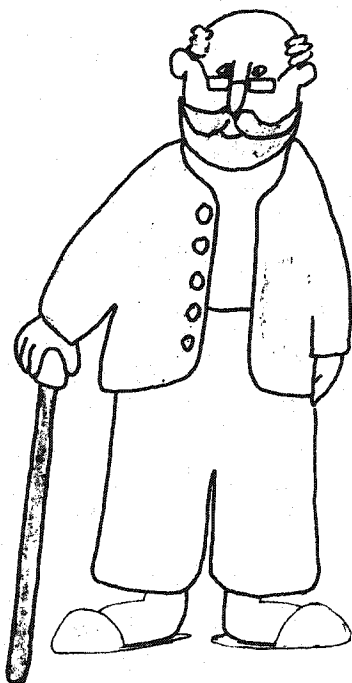


(el) milipo

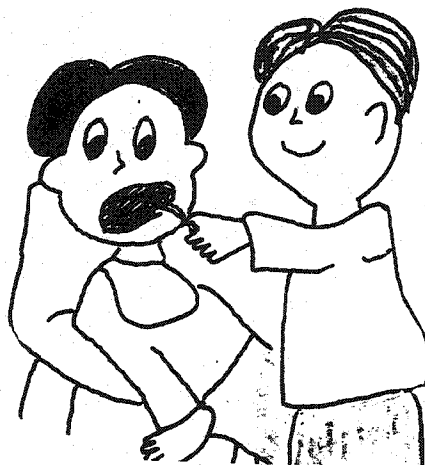


(la) capola

Real Words



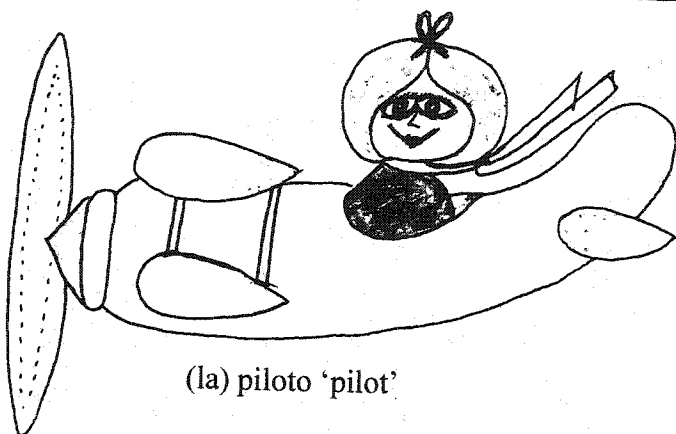
(el) abuelo 'grandfather'



(el) dentista 'dentist'



(la) enfermera 'nurse'



(la) piloto 'pilot'

Appendix B

Proportion of correct gender assignment per child, by condition.

	Natural	Linguistic	Conflicting	Converging	Sex	Year	ID
Right	3	3	3	3	B	1	18*
Wrong	3	3	3	3			
	0.5	0.5	0.5	0.5			
Right	1	4	1	4	B	1	15
Wrong	5	2	5	2			
	0.167	0.667	0.167	0.667			
Right	2	4	2	3	B	1	12
Wrong	4	2	4	3			
	0.333	0.667	0.333	0.5			
Right	4	2	5	3	B	1	8
Wrong	2	4	1	3			
	0.667	0.333	0.833	0.5			
Right	2	2	2	3	B	1	4
Wrong	4	4	4	3			
	0.333	0.333	0.333	0.5			
Right	1	4	4	6	B	1	38
Wrong	5	2	2	0			
	0.167	0.667	0.667	1			
Right	4	0	3	3	B	1	3
Wrong	2	6	3	3			
	0.667	0	0.5	0.5			
Right	2	5	3	3	B	1	36
Wrong	4	1	3	3			
	0.333	0.833	0.5	0.5			
Right	3	3	3	3	G	1	39*
Wrong	3	3	3	3			
	0.5	0.5	0.5	0.5			
Right	2	3	5	3	G	1	40
Wrong	4	3	1	3			
	0.333	0.5	0.833	0.5			
Right	6	4	4	6	G	1	29*
Wrong	0	2	2	0			
	1	0.667	0.67	1			

Right = number right in that condition
 Wrong = number wrong in that condition
 # = ID of participant
 * = participant dropped from analysis

B = boy
 G = girl
 K = kindergartener
 1 = first grader

	Natural	Linguistic	Conflicting	Converging	Sex	Year	ID
Right	4	4	3	3	G	1	22
Wrong	2	2	3	3			
	0.667	0.667	0.5	0.5			
Right	4	3	2	4	G	1	21
Wrong	2	3	4	2			
	0.667	0.5	0.333	0.667			
Right	4	1	2	3	G	1	17
Wrong	2	5	4	3			
	0.667	0.167	0.333	0.5			
Right	5	4	5	6	G	1	11
Wrong	1	2	1	0			
	0.833	0.667	0.833	1			
Right	3	3	2	1	G	1	16
Wrong	3	3	4	5			
	0.5	0.5	0.333	0.167			
Right	1	3	3	2	G	1	7
Wrong	5	3	3	4			
	0.167	0.5	0.5	0.333			
Right	4	6	1	2	B	K	10
Wrong	2	0	5	4			
	0.667	1	0.167	0.333			
Right	1	4	1	3	B	K	14
Wrong	5	2	5	3			
	0.167	0.667	0.167	0.5			
Right	4	2	2	5	B	K	24
Wrong	2	4	4	1			
	0.667	0.333	0.333	0.833			
Right	4	4	1	4	B	K	25
Wrong	2	2	5	2			
	0.667	0.667	0.167	0.667			
Right	3	5	3	5	B	K	28
Wrong	3	1	3	1			
	0.5	0.833	0.5	0.833			
Right	2	2	5	4	B	K	33
Wrong	4	4	1	2			
	0.333	0.333	0.833	0.667			

Right = number right in that condition
 Wrong = number wrong in that condition
 # = ID of participant
 * = participant dropped from analysis

B = boy
 G = girl
 K = kindergartener
 1 = first grader

	Natural	Linguistic	Conflicting	Converging	Sex	Year	ID
Right	3	3	3	3	B	K	31*
Wrong	3	3	3	3			
	0.5	0.5	0.5	0.5			
Right	5	4	4	1	B	K	30*
Wrong	1	2	2	5			
	0.833	0.667	0.667	0.167			
Right	1	2	2	5	B	K	42
Wrong	5	4	4	1			
	0.167	0.333	0.333	0.833			
Right	2	1	2	4	B	K	41
Wrong	4	5	4	2			
	0.333	0.167	0.333	0.667			
Right	4	3	4	4	G	K	1
Wrong	2	3	2	2			
	0.667	0.5	0.667	0.667			
Right	3	5	2	6	G	K	2
Wrong	3	1	4	0			
	0.5	0.833	0.333	1			
Right	2	4	1	2	G	K	5
Wrong	4	2	5	4			
	0.333	0.667	0.167	0.333			
Right	2	4	2	5	G	K	9
Wrong	4	2	4	1			
	0.333	0.667	0.333	0.833			
Right	4	2	2	6	G	K	6
Wrong	2	4	4	0			
	0.667	0.333	0.333	1			
Right	3	3	3	3	G	K	13*
Wrong	3	3	3	3			
	0.5	0.5	0.5	0.5			
Right	2	2	4	5	G	K	19
Wrong	4	4	2	1			
	0.333	0.333	0.666	0.833			
Right	2	5	4	3	G	K	20
Wrong	4	1	2	3			
	0.333	0.833	0.667	0.5			

Right = number right in that condition
 Wrong = number wrong in that condition
 # = ID of participant
 * = participant dropped from analysis

B = boy
 G = girl
 K = kindergartener
 1 = first grader

	Natural	Linguistic	Conflicting	Converging	Sex	Year	ID
Right	2	5	4	3	G	K	20
Wrong	4	1	2	3			
	0.333	0.833	0.667	0.5			
Right	4	3	4	1	G	K	23
Wrong	2	3	2	5			
	0.667	0.5	0.667	0.167			
Right	5	4	3	6	G	K	32
Wrong	1	2	3	0			
	0.833	0.667	0.5	1			
Right	4	2	3	3	G	K	26
Wrong	2	4	3	3			
	0.667	0.333	0.5	0.5			
Right	3	3	3	3	G	K	27*
Wrong	3	3	3	3			
	0.5	0.5	0.5	0.5			
Right	4	4	0	4	G	K	37
Wrong	2	2	6	2			
	0.667	0.667	0	0.667			
Right	3	6	5	5	G	K	34
Wrong	3	0	1	1			
	0.5	1	0.833	0.833			

Right = number right in that condition
 Wrong = number wrong in that condition
 # = ID of participant
 * = participant dropped from analysis

B = boy
 G = girl
 K = kindergartener
 1 = first grader

Appendix C

Results of paired 2-sample t-tests for means. All p values are two-tailed. Significant results are shown in bold.

t-tests for all children (n=34)

<i>Natural</i>	<i>Linguistic</i>	<i>Conflict</i>	<i>Converge</i>	t Stat	p(T<=t)
0.485	0.549			-1.141	0.262
0.485		0.456		0.577	0.568
0.485			0.632	-2.877	0.007
	0.549	0.456		1.591	0.121
	0.549		0.632	-1.526	0.137
		0.456	0.632	-3.353	0.002

t-test for kindergartners (n=20)

<i>Natural</i>	<i>Linguistic</i>	<i>Conflict</i>	<i>Converge</i>	t Stat	p(T<=t)
0.500	0.583			-1.269	0.220
0.500		0.425		1.072	0.297
0.500			0.683	-2.773	0.012
	0.583	0.425		1.986	0.062
	0.583		0.683	-1.255	0.225
		0.425	0.683	-3.587	0.002

t-tests for first graders (n=14)

<i>Natural</i>	<i>Linguistic</i>	<i>Conflict</i>	<i>Converge</i>	t Stat	p(T<=t)
0.464	0.500			-0.355	0.728
0.464		0.500		-0.493	0.630
0.464			0.560	-1.170	0.263
	0.500	0.500		0.000	1.000
	0.500		0.560	-0.836	0.418
		0.500	0.560	-0.891	0.389

t-tests for boys (n=15)

<i>Natural</i>	<i>Linguistic</i>	<i>Conflict</i>	<i>Converge</i>	t Stat	p(T<=t)
0.411	0.522			-1.177	0.259
0.411		0.411		0.000	1.000
0.411			0.633	-2.751	0.016
	0.522	0.411		1.046	0.313
	0.522		0.633	-1.234	0.238
		0.411	0.633	-3.347	0.005

t-tests for girls (n=19)

<i>Natural</i>	<i>Linguistic</i>	<i>Conflict</i>	<i>Converge</i>	t Stat	p(T<=t)
0.544	0.570			-0.388	0.702
0.544		0.491		0.753	0.461
0.544			0.632	-1.366	0.189
	0.570	0.491		1.206	0.243
	0.570		0.632	-0.891	0.384
		0.491	0.632	-1.780	0.092

Vita

Donna Bosworth Andrews is a resident of Seattle, Washington. She earned a Bachelors Degree in History and Religious Studies in 1985 from Lewis and Clark College (Portland, Oregon), and a Masters of Arts in Applied Linguistics in 1991 from the University of Southern California. She taught English as a second and foreign language for more than 10 years, living for extended periods in Europe and Asia. She has also taught Hellenistic Greek, Introductory Linguistics, and Second Language Acquisition. She is now a Software Usability Engineer and has published an article on conducting semi-structured interviews.