1 Introduction

1.1 Bilingual language acquisition

The terms *bilingual* and *bilingualism* have received diverse definitions. In this book, *bilingual* (the person), and *bilingualism* (the condition or state of affairs) refer to the use of two (or more) languages in everyday life.[1] Two major patterns of language acquisition have been identified in studies of early bilingualism: simultaneous bilingualism and sequential bilingualism, but no agreement exists with respect to the age at which bilingual development would be considered to be sequential. In simultaneous bilingualism, the child acquires two languages at the same time from birth or, as some researchers propose, before 3 years of age. Here, I use the term *Bilingual First Language Acquisition* (BFLA, or 2L1) to refer to situations where the child's exposure to two languages begins at birth (cf. De Houwer 2009: Ch. 1). This means that the question of the effect that different ages of first exposure to a language may have on the development of bilingual competence is not relevant in BFLA, but it is in sequential bilingualism. The latter could be differentiated, depending on when acquisition of a second language begins, into: (a) successive bilingualism, when the child's exposure to a second language starts sometime between the first and third birthdays; and (b) early second language acquisition, a form of early bilingualism that happens when a child has one established language before starting to hear and learn a second language (De Houwer 2009: 4).[2] This book focuses on BFLA – that is, on the acquisition of two languages from birth, Spanish and English in this case. The overall goal is to examine whether bilingualism affects the course of development in each language, and if so, how.

It has been estimated that half of the world's population, if not more, is bilingual (Grosjean 2010: 13; Tucker 1998). In the Spanish-speaking world bilingualism is indeed present, in all social classes and in all age groups, in every country where this language is official or co-official. Examples of languages

[1] The term *multilingualism* is usually used to refer to the use of three or more languages by an individual.
[2] See De Houwer (2009) for a comprehensive treatment of BFLA.

that share their communicative space with Spanish include: Mayan languages in Guatemala and parts of Mexico; Guarani in Paraguay; Quechua mainly in Peru, Bolivia, and Ecuador; Welsh in Patagonia (Argentina); Mapuche in parts of Chile and Argentina; Galician in Spain. Yet despite the situation of widespread bilingualism worldwide, which would appear to constitute typical BFLA learning contexts, research interest in the bilingual acquisition of Spanish with another language is rather recent.

The first in-depth longitudinal study of a bilingual child whose languages were Spanish and English was reported by Fantini (1985).[3] This linguist recorded the Spanish language development of his son from birth to age 10. Before him, two well-known studies not involving Spanish had been published: one by Ronjat (1913), of his son's development in French and German to the age of 4;10;[4] and the other by Leopold (1939–1949), who recorded the bilingual development of his two daughters in German and English. Until the 1980s, research on child bilingualism was scant, and in the USA largely limited to psychometric studies of school tests obtained from bilingual children. Since then, there has been an enormous growth of interest in the process of BFLA, perhaps made possible by advances in theoretical models of language and language acquisition (Hamers and Blanc 2000), coupled with the realization that child bilingualism is indeed prevalent throughout the world.

In recent decades we have started to ask the fundamental question of **how** infants acquire two languages simultaneously. Studies have shown that contrary to popular belief, being exposed to two languages from early childhood does not create confusion, but instead results in cognitive and social advantages (Bhatia and Ritchie 1999: 576). But many questions remain and continue to motivate studies about the development and effects of bilingualism: for instance, what is the relationship between social factors and different forms of bilingual behavior? What amount and quality of input is necessary for the development of bilingualism? Is language acquisition the same for bilingual and monolingual children? How are the two language systems represented in the mind of the developing bilingual child? Do the two languages develop independently or do they interact? What is the most compelling model to characterize bilingual language comprehension and production? The ensuing chapters address some of these questions.

[3] Fantini's 1974 doctoral dissertation reports his son's language development to age five. This is incorporated in the 1985 publication. Although Fantini's son heard Spanish and English from birth, he may not be the typical BFLA child. Indeed, his parents spoke English to each other, but did not address the child in this language, nor did they expect him to speak English in his infant years.

[4] Throughout the book, age notation follows the convention of separating years and months with a semicolon, and months and days with a period (e.g., 2;7.27 stands for 2 years, 7 months, 27 days).

Indeed, a crucial motivation for the study of BFLA is the fact that the simultaneous acquisition of two languages may allow us to see more clearly what principles are at work in the process of language acquisition (De Houwer 1990). Bilingual children are in fact the ideal subjects for crosslinguistic research because such factors as personality, cognitive development, and social environment are under control and not confounded as they might be in studies of monolinguals in one language or in language comparisons. In the case of bilinguals, different developmental paths could not be due to different stages of cognitive development, but rather to differences either in the languages involved or in the sociolinguistic conditions surrounding the two languages (i.e., the ecology of the languages). Thus, scholars of bilingualism have been attempting to answer a large number of questions that are also relevant to general theories of acquisition (Genesee 2000) and, beyond acquisition, to the study of language contact phenomena and language change more generally (Yip and Matthews 2007).

In this book, I describe and attempt to explain how two siblings, Nico and Brennan, developed the grammars and some lexical aspects of their two languages, Spanish and English, and how they achieved communicative skills in them. The siblings' language production was studied since they were about 1 year old until approximately 6 years of age. They were born to a two-language family in Los Angeles, were exposed to both languages from birth, and were expected to develop bilingually. There are no data about the number of children who grow up in dual-language homes around the world, but if my own family is an indication, these homes are nowadays far from being rare: I am Hispanic and my native language is Spanish; my spouse is an English-speaking Euro-American; our three sons are Spanish-English bilinguals, two are also married to English-speaking Euro-Americans, and one is married to a bilingual Philippine-American. On any given day when the family gets together, Spanish, English, and Tagalog may be heard in our home.

Comprehensive case studies of bilingual first language acquisition have examined children from their first word up to about age 5;0 (e.g., De Houwer 1990; Deuchar and Quay 2000; Lanza 2004; Meisel 1990, 1994; Yip and Matthews 2007). These studies have been concerned with the examination of the linguistic and cognitive skills that underlie bilingual children's language capacity. Though none have studied in depth the Spanish-English bilingual development of two siblings from their first word to age 6, as this book does, I follow their analytical tradition and that of others who have shaped the field. A comparison with their findings will surely strengthen some of the outcomes of my study, which include the following:

- The morphosyntactic structures of the languages of a bilingual are acquired in ways that are not radically different from monolingual acquisition: for each of their languages, bilingual children use similar structures at similar

stages in development and make similar types of errors as their monolingual peers.

- Differences with monolingual acquisition are accounted for in ways which allow for anticipating what language aspects may be affected by crosslinguistic interaction and delayed acquisition, and under what conditions. A key factor determining the differences observed concerns the social properties of Nico's and Brennan's language input (reflected, for instance, in quantity of input), and the quality of the input (e.g., regularity, frequency, and functional transparency of language structures).
- As developing bilinguals grow older and receive more input in one of their languages and have more opportunities for its use, this language becomes stronger and motivates the transfer of some its features to the weaker, less-used language. Changes that distance the siblings from the grammars of the L1 adult models become noticeable from about the age of 4;0.
- Different language domains are affected differently by unbalanced exposure to and use of two languages.
- All other things being equal, the birth order of the child within the family accurately predicts the level of bilingual proficiency achieved. As observed in family after family, the older child receives a larger amount of direct input and achieves a higher level of bilingual proficiency than the younger siblings.

The studies presented in the ensuing chapters examine naturalistic data from two developing bilingual siblings who have each received different amounts of input in the languages at different stages through their first six years, and also different amounts of input compared to each other. This unequal amount of input is shown to have an effect on the level of productive proficiency achieved by the children in their weaker language, Spanish. A relatively reduced exposure to this language correlates with a lower level of productive proficiency and greater distance from the input provided by the adults in the children's environment.

The practical implications of the present study are many. It provides an explicit description of key aspects of the emerging grammars of the bilingual siblings. In the study of tense acquisition, the book also examines the development of the early verb lexicon and includes a comparison with two Spanish-English bilingual preschoolers with diverse sociolinguistic histories. The results will contribute to our growing knowledge of the processes involved in bilingual acquisition and the factors that underlie individual differences in bilingual development. The outcome of this research may play a part in the resolution of educational questions, and may also bear on important theoretical issues related to such processes as transfer, language change, and the role of intra- and crosslinguistic factors in determining degree of bilingual achievement. I also suggest answers to the question of the mechanisms that account for cross-linguistic interactions. This is related to the issue of what makes linguistic

systems permeable to external pressures and may contribute to the construction of theories about the grammars of languages in contact.

1.2 Theoretical preliminaries

A discussion of the many theories that have been proposed to explain how children go about the process of acquiring language would require more than a section in a book – indeed more than a book (see Ambridge and Lieven 2011 for an evaluation of the major current theories).

Studies of BFLA have been done within the framework of the theories proposed for monolingual acquisition, in particular, *nativist* and non-nativist *constructivist* theories. The latter have included functionalist, emergentist, socio-pragmatic, usage-based proposals, and the competition model.

Nativists view language as a fundamental part of the human genome. They argue that children are genetically endowed with a Universal Grammar (UG), a set of linguistic principles and constraints common to all the world's languages, which, together with language-specific parameters, guide acquisition once children are exposed to a particular language. Language is not conceptualized as a cultural object, but as an instinct whose development is a natural part of maturation, no different from dolphins learning to swim or spiders learning to spin spider webs (Pinker 2000). In the nativist generative view, the information obtained from the input is insufficient for children to learn a language; the input is frequently messy and unstructured, and does not provide negative evidence, so some aspects of their linguistic knowledge must be innate. The argument of the poverty of the stimulus has been called into question, however, based on evidence showing that children receive sufficient input from caregivers and peers, and that this input is clear, ritualized, and embedded in contexts that make language comprehensible to the child and help her grasp the way in which language is structured and used (Tomasello 2003: Ch. 8, 2005). A further challenge to the theory of innateness has been the identification of what constraints or structural features are due to nature (rather than nurture) – that is, the question of what is hard-wired in the mind remains largely unanswered.

In addition to innate knowledge, nativists propose that children have available to them two sets of learning procedures, one set of domain-general procedures that are used to acquire words and exceptions to general rules, and a set of language-specific learning mechanisms for those aspects of language that conform to UG and core regular rules, such as the formation of regular past tense and the order verb-direct object in English. This tenet has also been contested by scholars who argue that the learning of words and exceptional or regular grammatical structures is achieved with the same set of learning mechanisms that are a part of a general cognitive learning apparatus. These domain-general mechanisms are used

for forms of learning that are linguistic and non-linguistic (Bates and Goodman 1999).

A great deal of thought-provoking research on bilingual language acquisition (BFLA, adult L2 acquisition, etc.) has been conducted within the nativist generativist model. In particular, work on the interfaces between language modules (e.g., syntax–semantics interface, syntax–discourse interface) has been intense and fruitful (see, among others, Hulk and Müller 2000; Montrul 2011; Serratrice, Sorace, and Paoli 2004). I refer to some of this work in the ensuing chapters.

Usage-based constructivist theories assume that language emerges from the interaction of cognition with experience. What is innate is general cognition and the ability to learn. From birth, the child is endowed with cognitive abilities and general learning mechanisms that make possible the learning of language from the input received in situated instances of communication – that is, in instances of social interaction. Language is learned like other cognitive skills (e.g., how to read, solve puzzles, organize materials) by pragmatic inference, analogy making, imitation, sensitivity to frequencies in the input,[5] and by social interaction with caregivers and others.

The usage-based model proposes that the child may be endowed with innate semantic and pragmatic notions, but not with morphosyntactic categories, rules, or representations. Rather, syntactic and morphological categories emerge from concrete pieces of language which illustrate such concepts as object, place, event, agent, and gradually develop into more schematic constructions. Linguistic constructions are conceptualized as complex linguistic symbols with a form, a meaning, and a pragmatic function.[6] Constructivism assumes a bottom-up view of morphosyntactic acquisition, similar to the acquisition of the lexicon (vocabulary): from instances to the general rule. Children are viewed as pattern builders, born equipped with powerful cognitive and pragmatic abilities that allow them to induce patterns from the particular, language-specific utterances to which they are exposed (Tomasello 2005).

Usage-based approaches have shown that morphosyntax emerges slowly and in a piecemeal fashion, varying between target and non-target forms: learning is item-based; no abrupt, all-encompassing learning of a grammatical aspect occurs; there is always a period of variation (Gathercole, Sebastián, and Soto 1999). Abstract rules of language develop gradually from the characteristics of

[5] Henry (2004), for instance, cites her study of English monolinguals aged 2;0–4;0 years, in which she shows that children not only acquire variable forms of a structure at an early stage, but also use them "in frequencies similar to those occurring in the input" (p. 279) to which they are exposed. This strongly supports the view that acquisition must be frequency-sensitive.

[6] There are also constructivist proposals that do not make reference to meaning or social-pragmatic understanding: for instance, distributional evidence available in naturalistic speech may offer strong cues to grammatical form-class assignment (Ambridge and Lieven 2011; Weisleder and Waxman 2010).

the natural language input children receive once a critical mass of exemplars has been reached (Akhtar and Tomasello 1997; Marchman and Bates 1994). The challenge to this theory remains the identification of what constitutes a "critical mass of exemplars."

A tenet for functional, usage-based proposals is that communication is crucial for the child's survival and therefore the acquisition of language is guided by the child's need to use language to perform communicative functions (e.g., to request an object or activity, to influence the actions of others, to enquire about entities) and to understand what others say. Language cannot be explained, therefore, without reference to its communicative functions (Ambridge and Lieven 2011: 3).

There is an important body of research that provides support for some of the principles of usage-based theories of monolingual language acquisition: the key role of analogy and the relative type and token frequencies within and across constructions; an item-based approach to learning; "starting small" in acquisition; children's ability to read a speaker's focus of attention and communicative intentions; their extraordinary ability to imitate the things that other people do (see, among others, Goldberg, Casenhiser, and Sethuraman 2004; Kidd, Lieven, and Tomasello 2006; several contributions in Rojas Nieto 2003, 2009; Tomasello 1998). Because these theories regard meaning and the social situation in which language is normally produced as key to the understanding of how language is acquired, they provide a solid theoretical foundation for the study and explanation of how a child learns two languages at a time – that is, as first languages (2L1).[7]

This is not to say that usage-based theories (or nativist approaches) have provided the definitive answer to how language is acquired. Indeed, Tomasello (2003), a staunch defender of usage-based theory, who has contributed much to its development, acknowledges that there are no fully adequate theoretical accounts of how young children go about the business of developing language. Which of these theories more appropriately accounts for how children learn language is an open question that will surely continue to motivate many studies and exciting debates.

Given the state of our knowledge, I favor a theoretically eclectic empirical approach to the investigation of bilingual development. Although mostly informed by usage-based and variationist principles, the studies I have conducted also rely on the assumptions of language contact theories (as discussed in, among others, Matras 2009; Prince 1992; Silva-Corvalán 1994; Thomason 2001; Thomason and Kaufman 1988; Winford 2003), and incorporate an interest in reviewing work that examines the role that interfaces between

[7] Indeed, this theoretical perspective has been implicit in some of the most influential work on BFLA, such as that of De Houwer (1990), Lanza (2004), Leopold (1939–1949), and Ronjat (1913).

language components play in determining vulnerability to influence across languages. There is a natural harmony between usage-based and variationist approaches: both highlight the need to examine language in its natural contexts of use, and the value of quantification and empirical accountability (the "principle of accountability": Labov 1972: 72). Accordingly, qualitative and quantitative analyses of a wealth of longitudinal data collected from the two siblings are included in this book. For example, in order to describe and explain the constraints on pronominal subject expression in Spanish and English, and the ways in which the siblings' grammars differ in this respect, I draw extensively on discourse analysis (functionalism) and on variationist techniques of quantification. Explanations of the patterns of Spanish subject use identified are based on the prediction made in language contact theory to the effect that discourse-pragmatic restrictions are vulnerable in a situation of intensive bilingualism (Prince 1992; Silva-Corvalán 2008). This outcome is also related to interface theories that make similar predictions (e.g., Müller and Hulk 2001).

1.3 Models and issues in bilingual first language acquisition

1.3.1 Language differentiation models

The study of bilingual development presents specific theoretical questions beyond those that nativists and non-nativists engage in. The fundamental goal that research on BFLA (or 2L1) has set up is the development of models that may account for how bilinguals acquire two linguistic systems at a time, how these systems interact, and how they are represented in the mind of the bilingual. Some researchers have observed that the principles put forward by constructivists, for instance, are also valid to account for the acquisition of 2L1 (e.g., Gathercole 2007). In addition, 2L1 researchers have proposed models and questions specific to bilingual acquisition. One important debate has centered around the question of whether bilinguals begin acquisition with two independent systems or with a fused one. Two opposing models have been proposed: a *unitary language system* model (ULS) versus a *separate development* model (SDM), and a related model of *autonomy or differentiation with interdependence* of the systems of the bilingual.

The ULS, defended by Volterra and Taeschner (1978), asserts that although children are exposed to distinct sets of linguistic input, they go through an initial stage in which their two or more languages are represented in a unitary or fused system until they reach a stage of differentiation at around age 3;0 (see also Murrell 1966; Schnitzer and Krasinski 1996). Initially, the child has one lexicon with words from both languages (no equivalent words in the two languages). At a second stage, the child may differentiate the lexical systems, but produces

syntax reflecting only one system. Finally, the child reaches a stage at which syntax and lexicon are differentiated.

This "early confusion" view of bilingual acquisition is no longer supported, however, at least not for morphosyntactic development. Exponents of the SDM maintain that from the earliest appearance of phonology, morphology, and syntax, forms are used in a language-specific manner (Montanari 2010). When a child is regularly exposed to two languages from birth and is acquiring a fairly balanced level of proficiency in both, each language system develops in a self-sufficient and independent manner (De Houwer 1995, 2005; Deuchar and Quay 2000; Meisel 1989, 2001; Paradis and Genesee 1996). This book provides evidence in support of this view, although there also seems to be some indication of a stage in very early acquisition when it is difficult to determine that some aspects of the two languages, English and Spanish in our case, are indeed clearly differentiated. I will argue, however, that these instances of non-differentiation may be due to similarities between these two languages, coupled with features of early language development.

In the search for confirmation of the separate systems model, numerous studies have shown that bilingual children follow a path of development that is comparable to that of monolinguals in each of their languages, and produce mostly language-specific structures from the start of bilingual production. But BFLA research has also shown that simultaneous bilinguals also produce some unusual structures not attested in the monolingual acquisition of the relevant languages nor in the input they receive from caregivers. This outcome has been interpreted to be proof of some degree of crosslinguistic interaction or transfer of features between the languages being acquired by the child. Thus, the model of how BFLA children acquire their two languages that appears to have the clearest support from current research is one of *autonomy with interdependence*. This model assumes, then, differentiated development with some degree (to be empirically determined) of crosslanguage interaction – that is, interdependence. The interesting question is no longer whether the two languages are acquired in a separate fashion, but rather what it is that may be affected when two languages are acquired from birth, and why. This interest has motivated much enlightening research, while at the same time raising questions about theoretical concepts in need of clarification. In the following paragraphs I discuss some of these concepts and deal with them again in the various chapters.

Central to the discussion is the issue of what crosslinguistic interaction is, frequently discussed in the literature as coextensive with crosslinguistic influence, transfer, and at times also with the term *interference*, which I prefer to avoid because of its negative connotation. Here, I discuss ways to define crosslinguistic interaction, the factors that favor it, and its various manifestations. This discussion requires clarification not only of the concepts of crosslinguistic interaction, influence, and transfer, but also of balanced and

unbalanced bilingualism, dominance, proficiency, incomplete acquisition, and system instability.

I do not review here the many ways in which these notions have been said to play a role in models of bilingual acquisition and crosslinguistic interaction, but make an effort to define them and apply them consistently in accounting for the BFLA phenomena to be presented in the ensuing chapters.

1.3.2 Defining crosslinguistic phenomena

As the term indicates, crosslinguistic interaction refers to the effect that the lexicon and grammar of one language has on the lexicon and grammar of the other language spoken by a (developing) bilingual.[8] This effect may be temporary, sustained over a period of time, or permanent, and it may be manifested in several different ways. Thus, one of the languages may affect the development of the other language by delaying or accelerating acquisition, by causing the production of non-target constructions not attested in monolingual development, by promoting the use of superficially parallel constructions in inappropriate discourse-pragmatic contexts, or by transferring forms and meanings. I propose the term *crosslinguistic interaction* as an umbrella term encompassing any type of effect that results from the coexistence of two languages in the mind of the bilingual.

Delayed acquisition of a grammatical phenomenon present in one of the languages is established if, compared with monolingual acquisition, the language feature is acquired at a later age (e.g., delay in the acquisition of obligatory subjects in English by English-Spanish bilinguals). The opposite, earlier acquisition by bilinguals compared with monolinguals, would be acceleration of acquisition, which could occur when a grammatical feature is marked similarly in both languages (e.g., plural marking in Spanish and English). The production of non-target constructions is frequently the result of the transfer of a linguistic feature from one language to the other – for example, the use of stranded prepositions in the Spanish of English-Spanish bilinguals (*¿Qué abres la lata con?* 'What do you open the can with?'). Transfer from Spanish or French to English could result in the use of, for example, post-nominal adjectives (*a book red*, cf. Nicoladis 2006). In the data examined here, the children go through a stage of syntactic transfer and produce stranded prepositions in Spanish, but these structures are later discarded. I find this comparable to the stage monolinguals go through when they experiment with such constructions as "We goed

[8] The influence may also involve the phonetics and phonology of a language, but this book considers phonetic aspects only when required by the discussion of a lexical or grammatical phenomenon.