Patterns of Interaction in the Lexical Growth in Two Languages of Bilingual Infants and Toddlers

Barbara Zurer Pearson  Sylvia C. Fernández
University of Miami  University of Maryland

We investigated the extent to which bilingual children follow the same patterns and timetable of lexical development as monolinguals. For a group of 20 simultaneous bilingual (English-Spanish) infants, ages 10 to 30 months, we looked at the patterns of growth in one language in relation to growth in the other and also with respect to growth in both languages combined. The MacArthur Communicative Development Inventories (CDI), standardized parent report forms in Spanish and English, provided measures of lexical growth in two languages at varying intervals within the age range. We plotted the two single-

Portions of this paper were presented at the Southeast Conference on Linguistics' annual meeting, April 1992, Gainesville, Florida, and at the Eighth International Conference on Infant Studies, May 1992, Miami Beach, Florida. This research was supported in part by Grant 5 R01 DC00484, NIH NIDCD and IR01 HD30762/01 NNIE-SBIR. We gratefully acknowledge the help of Donna Thal and Donna Jackson-Maldonado of the Center for Research in Language at University of California, San Diego for the use of the Spanish forms. We give special thanks to the Family Coordinators, Debra Basinger, Vanessa Lewedag, and Sharyse Levine, and to the families who participated and persevered.

Correspondence concerning this article should be addressed to Barbara Zurer Pearson, Bilingualism Study Group, University of Miami, Department of English, Post Office Box 248145, Coral Gables, FL 33124-4632. E-mail: bpearson@umiami.edu
language measures, as well as Total and Total Conceptual language measures, across time, referenced on a second y-axis to the percent of the child's language environment that each language represented. For a subset of the children, we calculated the percentages of general nominals, social words, and verbs for each language to allow the characterization of the children's learning strategies as "referential" or "expressive" (Nelson, 1973). The rate and pace of development were similar to patterns observed in monolinguals. In addition, the vocabulary spurt was seen to occur in about the same percentage of children as has been observed in groups of monolingual children. The bilinguals differed from one another with respect to the relative independence of one language from the other, including the use of different learning strategies in the two languages by the same child.

A substantial minority of bilinguals experience what has been called "bilingualism as a first language" (Swain, 1972). In surveys of Hispanic populations in Miami (Pearson, 1993; Pearson & McGee, 1993), for example, between 6% and 15% of respondents report having learned their two languages from birth. As half of the 33,000 children a year born in Miami are of Hispanic background (Florida Office of Vital Statistics, 1993), even 10% would amount to about 1,600 new simultaneous bilingual children per year in Miami alone and the number is growing faster yearly (Merzer, 1993). However, the extent to which monolingual models of language acquisition adequately explain bilingual development has not been widely explored. Some studies have compared bilinguals' development in one of their languages to monolinguals', but little research has examined simultaneous growth in two languages and possible interactions between them (McLaughlin, 1984). This question is of special scientific interest for linguistics (Levy, 1985; Meisel, 1989; Snow, 1988) and of practical interest for educational psychology. A detailed description of the course of simultaneous bilingual acquisition is necessary to achieve a fuller understanding of the special capabilities and needs of bilingual children.
However, generalizations about bilingual development will be even more difficult to make than those about monolingual development. A survey of bilingual households in Miami, undertaken for our study, reveals a wide variety of patterns of language use in the homes where bilingual children are being raised. It will not be surprising therefore to observe several general patterns of development, reflecting the influence of these varied circumstances on an already wide range of individual differences observed in typically developing monolinguals.

Until now, little descriptive information has been available about early bilingual acquisition. Careful documentation of early lexical development, for example, has been reserved almost exclusively for the children of linguists. Just a few published and unpublished word-lists, derived from case studies, represent the vocabulary at various ages for fewer than a dozen bilingual infants (Jekat, 1985, as cited in De Houwer, 1990; Leopold, 1939; Mikes, 1990; Quay, 1993a, 1993b; Taeschner, 1983; Vihman, 1985; Vogel, 1975; Volterra & Taeschner, 1978; Yavas, 1991; see De Houwer, 1990, for a review). Without detailed information about a wide range of children growing up in varied circumstances, it has been difficult to gain a broad perspective on what is typical and what is exceptional in early bilingual development. It is crucial, therefore, to move beyond the case study method in bilingual studies. To broaden the base of observation, it is important to examine a restricted question, from a consistent point of view for many children.

Fortunately, new diagnostic and research tools have facilitated observation of early bilingual lexical development. The development of standardized inventory forms for vocabulary from ages 8 to 30 months (MacArthur Communicative Development Inventory, 1989; Dale, Bates, Reznick, & Morisset, 1989) has allowed parents without special training to make a scientifically useful record of their children's earliest words. Such inventories, which are becoming available in several languages, including English and Spanish, create a valuable new data source about the patterns of early lexical development. In our study, comparisons
for individual children of two single-language inventories at a number of time points give a graphic picture of the stages of lexical growth in each language, separately and together.

We observed bilingual growth in part to compare it to the patterns of vocabulary development for monolingual children. In this age range, typical monolingual development is very diverse. Our general expectation in productive vocabulary (Dromi, 1987; Goldfield & Reznick, 1990; Nelson, 1973) is that children begin producing their first recognizable words around one year of age. The majority of children acquire first words rather slowly until a burst, between around 25 or 50 words, after which a sharply faster rate of acquisition, the so-called "lexical spurt," has been observed until about 50 or 100 words (Dromi, 1987; Goldfield & Reznick, 1990; McCarthy, 1954; Reznick & Goldfield, 1992). The Technical Manual for the MacArthur Communicative Development Inventories (CDI) by Fenson et al. (1991), which presents vocabulary measures for 1,600 children, corroborates these expectations from the earlier literature. For the CDI norming sample, the age of first words is generally between 9 and 14 months (Fenson et al., Tables C5 and C6); the median age for producing 50 words is between 16 and 17 months, with a range roughly from 12 to 25 months; and the median age for 100 words is between 18 and 20 months, with a range from 13 to 27 months (Fenson et al., Figure E6). Girls tend to learn words slightly earlier than boys—Fenson et al. chart their growth separately—but normal development for both genders is very diverse. Standard deviations around the mean number of words at a given age are greater than or equal to the mean through roughly 18 months (Fenson, 1991). The charts from the CDI norming sample do not give evidence detailing a lexical spurt, because different children experiencing a spurt at different times are averaged out. Nonetheless, their data provide a useful context within which to evaluate early lexical growth for individual children.

Another useful framework for the examination of lexical growth is the referential/expressive distinction first proposed by Nelson (1973). The extent to which children rely on general nominals (names for things, as opposed to names for individuals or
social words) in building vocabulary reveals what Nelson calls a strategy for learning and using language. Children with more than 50% general nominals in their vocabularies at a given point are said to use a "referential," largely "object-oriented" strategy, while those with fewer than 50% general nominals will have more "personal-social" and "specific nominal" words, indicating an "expressive" or "more self-oriented" interactional language style (p. 22).

Bilingual children’s use of these strategies in early acquisition is largely unexplored, but because these strategies are thought to arise as a result of the child’s cognitive preferences as well as pressures in one direction or the other in the environment (Goldfield, 1986), we thought it important to see how consistent such patterns are in one child’s behavior in two different languages and language environments. Bates, for example, reported that her daughter Julia, referential in her earlier English acquisition, appeared to adopt a more expressive, social style when exposed to Italian for the first time at about 20 months of age (Bates, Bretherton, & Snyder, 1988, chap. 16).

Finally, Marchman and Bates (1991) proposed that growth in the verb lexicon is a revealing marker of linguistic and conceptual maturity. Figure G7 of the CDI Technical Manual (Fenson et al., 1991) shows average verb production for their sample of 1,600 English-learning children to be 0% at vocabulary sizes of 50 words or less, 5% at 51–100 words, 10% at 201–300 words, and peaking around 17% at about 500 words. Similar percentages were reported for Spanish-learning children (Jackson-Maldonado, Thal, Marchman, Bates, & Gutiérrez-Clellen, 1993). Because conceptual development is often presumed to be independent of any specific language (Cummins & Swain, 1986), it will be of interest to explore whether verb growth in each language is tied more closely to growth in one language alone or to growth in both.

Goals of the Study

The extent to which bilingual children follow the same patterns and timetable of lexical development as monolinguals is still
an open empirical question. Pearson, Fernández, and Oller (1993) compared vocabulary sizes in production and comprehension in children between 8 and 30 months and found no statistically significant quantitative differences between a bilingual and a monolingual sample. The present research looks more closely at the qualitative patterns of growth of vocabulary knowledge in the same bilingual sample, with special attention to the growth in one language in relation to growth in the other and also with respect to growth in both languages combined. We address the following questions:

1. How similar are the patterns of bilingual lexical development in each language to the broad outlines of lexical development observed for monolingual children? Do bilingual children, like monolinguals, generally exhibit a lexical spurt after an initial period of slow steady growth?

2. How closely does development in one language reflect the development in the other? For example, how often do children achieve a balance between the languages? Are any interactions observable between the rate of learning in one language and the rate of learning in the other language? If a spurt is observed, is it observed in each language, or in only one? How independent is the vocabulary learned in one language from the vocabulary in the other? Do children have primarily singlet vocabulary, words known in only one or the other of the languages, or do they tend to have doublets, the same concept lexicalized in both languages?

3. Finally, in what ways can growth in one language be seen as distinct from growth in the other? Do children appear to use different learning strategies in their two languages? How does growth in the verb lexicon of each language appear to relate to general growth within and across languages?

The 20 children investigated here represent too small a basis for normative statements about bilingual development. Nonetheless, charting their experience enlarges considerably the scope of our knowledge about bilingual children's language development.
Methods

Participants

The current research on lexical development is part of a five-year study of infant vocalizations in different populations. Participants were recruited soon after birth, through health department records and word-of-mouth solicitation. Eighteen children being reared in English-Spanish bilingual homes, 8 females and 10 males, provided the longitudinal data for this study. The number of vocabulary observations varied from 2 to 10 per child at approximately 2- to 4-month intervals between the ages of 8 and 30 months. All but two children came from middle-class homes. The children were all of normal intelligence with an average Bayley (1969) score at 18 months of 113 (SD=12.9). Two children were approximately five weeks premature, with no other health problems. For the analysis of learning strategies, two single observations of middle-class, full-term children from the vocal development study were also included. In all, there were 77 observations, 72 of production and 5 of comprehension vocabulary.

All of the children had significant exposure on a regular basis to both English and Spanish through their various caretakers, who were native speakers of one or both languages. In some households, children of parents who were monolingual in one language had caretakers who were speakers of the other language. More often, one parent and his or her extended family were native speakers of one language and bilingual to varying degrees, and the other parent was a speaker of the other language. In still other households, both parents were bilingual, and they had various childcare arrangements, generally involving monolingual-Spanish grandmothers or nannies. Such arrangements could be consistent for the child over time, but in some cases the language profile changed with a family move, the addition of new members to the household, different work patterns for the parents, or changes in the childcare options available.

Although all parents expressed a desire to provide an envi-
vironment balanced equally between the languages, and the condi-
tions of their households appeared to support that desire, only one
child had approximately equal exposure to both languages during
the period of observation. Parent estimates of language exposure,
updated at regular intervals, averaged 60% to 65% of one language
and 35% to 40% of the other; 12% of the children had an exposure
less balanced than 75:25. Nine of the 18 children with more than
one observation experienced a relatively consistent language en-
vironment throughout the data collection period; 9 children
experienced changes in the percentage of time they were exposed
each language. These latter included 4 who experienced changes
in which language predominated. Of the other 14 children, 9 spent
more time in a Spanish environment, 4 spent more time in an
English environment, and the last child (as noted above) heard
equal amounts of both languages. Because of their probable
influence on the relative utility of the languages for the child,
changes in the language environment are noted here whenever
language measures are reported.

Materials

We used standardized parent report instruments, the
MacArthur Communicative Development Inventory (CDI), Tod-
dler and Infant forms (1989) and their Spanish adaptations, the
Toddler and Infant Inventario del Desarrollo de las Habilidades
Comunicativas (Jackson-Maldonado & Bates, 1988), to assess the
vocabularies of the children. The Infant English form contains 395
words frequently produced and understood by infants between 8
and 15 months, arranged in 22 semantic and grammatical catego-
ries. The Toddler form, for use between 15 and 30 months,
contains 679 words. Instructions on the Infant form tell parents
to mark in one column words their child has comprehended only
and in a separate column words the child has both comprehended
and spontaneously produced; on the Toddler form parents mark
only the words that their child has produced. The vocabulary
scores are the number of words marked by the parent, one number
for comprehension, another for production. All vocabulary reported here, unless otherwise noted, is production vocabulary.

The Spanish version of the CDI was developed by adopting the format of the English (and Italian and Japanese) Inventories, but using Spanish word-lists and research studies to dictate the items included. It lists 428 words on the Infant form and 732 words on the Toddler form. The version of the Spanish CDI used in our study was modified slightly to include lexical items used by the Cuban-American population of Miami (Fernández & Umbel, 1991).

The CDI is more effective than previous parent report measures because it relies on the parents’ recognizing, rather than recalling, the words in the child’s vocabulary. Additionally, the two forms of this inventory focus on “emerging behaviors” at times when these behaviors are current (not retrospective) and limited in number. Its upper limit at 30 months of age reflects the fact that most typically developing children produce so many words at that age that parents can no longer keep track of them. After 30 months, other assessment instruments are more appropriate. The CDI vocabulary list for a given child is not a true inventory, as the form does not exhaust the list of possible words children might say. Rather, like vocabulary tests at later ages, it requires an extrapolation of the total vocabulary based on a controlled sample. Evidence of the CDI’s reliability and validity is reported in Fenson et al. (1991)\(^1\). In the research reported here, the comparison of one language to the other using the same measure is more important than the precise number of words known.

Procedure

The children’s parents filled out language background questionnaires which they updated each time they filled out the CDIs. They estimated the amount of time per day or per week that the child spent with speakers of each language or, if the children were with bilingual speakers, what percentage of each language was spoken with them. The number of observation points per child depended on the length of time the family took part in the parent
grant and also on the parents' level of cooperation. In some cases, one individual filled out both language forms for the child; in others, one parent did the inventory for one language and the other parent (or caretaker) did it for the other language. Only three of the children had observations across the entire age range. Therefore, depending on the children's ages at the end of the study and their stages of development during their family's participation, different subsets of children, ranging from 7 to all 20, provided the relevant data to answer the various research questions and subquestions posed above.

In the instructions to parents, we emphasized that the vocabulary inventories were measures of spontaneous vocabulary production rather than of prompted repetition. As indicated in the CDI instructions, we told the parents to mark words that their child said even if the pronunciation was incorrect. Thus the consistent pairing of a certain sound with a particular meaning was sufficient for the parents to mark off that word even if their child's production of the word was different from the adult pronunciation. Wordforms which were used for more than one concept within a language, such as ba for ball and ba for baby were counted separately because they reflected two sound-meaning pairings.

The information provided by the CDI is only approximate however, in that the parent is not asked to specify the referent of a word. At these ages, children's meanings for words are often either overextended or underextended, as compared to the adult definition. Ball, for example, may overextend to anything round or anything one throws, whereas a word like zapatos may underextend to refer only to one particular pair of sneakers. No claims are made in the use of the CDI that the children's words have identical meanings to adults', just that they have begun to use them in ways their caretakers can respond to.

To analyze the relative vocabulary sizes of these bilingual children, four measures were taken: two measures, English and Spanish Vocabulary, were taken directly from the respective CDIs; two double-language measures, Total and Total Conceptual Vocabulary, were constructed from the comparison of the single-
language forms. The first two, the single-language measures, were like monolingual measures except that each bilingual subject had two "monolingual" assessments. The two double-language measures encompassed the child's abilities in both languages summed together and then corrected for knowledge shared between them (with the cautions expressed above). Words that shared the same phonetic shape across languages were counted in each single-language inventory, but only once in the measure of Total Vocabulary. (Phonetic similarity was estimated from our own experience listening to tapes of these children and after conversations with the mothers. Generally, the words considered to be phonetically similar translation pairs were animal sounds, some proper names, and a small set of cognate terms common in child speech like *mama* and *choo choo*, etc.). This measure is smaller than the combined English and Spanish totals, in that a single word-form that served the child in both languages was considered to be only one word. The fourth measure, Total Conceptual Vocabulary (TCV), was a count of the child's lexicalized concepts or linguistic functions, whether they were coded in one or both languages. In effect, the TCV included all the words in one language plus only the singletons, or words coded uniquely, in the other language. In this manner, we counted cross-language synonyms, such as *dog* and *perro*, as one word in Total Conceptual Vocabulary but within-language synonyms, like *hi* and *hello*, as two, (just as in the monolingual vocabulary calculation).

For these double-language measures, we needed to determine how many words the child had coded in only one language and how many were coded in both. To know when a word checked on one form was similarly checked on the other, one form had to be mapped onto the other to the extent possible (Pearson, 1992). The first step in the mapping process was to compare the English and Spanish versions of the CDI and determine which words could be termed Translation Equivalents, or "doublet pairs". For the most part, this was fairly straightforward. The English and Spanish words in each pair were both assigned a unique "pair number". For instance, *dog* and *perro* were both given the number 214, *table* and
were assigned number 927, and so forth. However, due to cultural and linguistic differences between Spanish and English and, in some cases, simple gaps on one form or the other, not all words could be paired. In all, about 80% of the words were matched with a translation equivalent on the other form. For the children examined here, words without a potential pairword on the other language form tended to be less commonly used, and so we were able to conduct an analysis of translation equivalents on an average of 88% of each child’s reported vocabulary.

As an example, consider a child with the following words:

<table>
<thead>
<tr>
<th>English</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>mama</td>
<td>mama</td>
</tr>
<tr>
<td>bear</td>
<td>oso</td>
</tr>
<tr>
<td>duck</td>
<td>abuela</td>
</tr>
<tr>
<td>more</td>
<td>agua</td>
</tr>
<tr>
<td>daddy</td>
<td>si</td>
</tr>
<tr>
<td>no</td>
<td>araña</td>
</tr>
</tbody>
</table>

The child would be considered to know 6 words in Spanish and 6 words in English, or 12 words total. If mama were reported by the mother to be pronounced the same in both languages, the child’s Total Vocabulary would be 11 words, and Total Conceptual Vocabulary would be 10 (6 Spanish, plus 4 nondoublets in English, or vice versa). Doublet words (not pairs) would be mama and bear—oso, equaling 3.

We plotted the four basic measures of vocabulary size—English, Spanish, Total, and Total Conceptual—across time for each child. These graphs allow us to compare growth in one language to growth in the other and to growth in both combined. The number of doublet pairs, represented by the relation between Total and Total Conceptual Vocabulary, provided the basis for estimating the degree of lexical overlap from one language to the other at different stages of development. In addition, we graphed the information about the children’s language environment on a second y-axis to evaluate when changes in the language measures could be directly attributed to changes in the environment.

Finally, for a subset of the children representing a range of
patterns of exposure to the two languages, we examined the
wordlists themselves to compare the types of words known in each
language. We calculated the percentages of "general nominals" for
each language to allow the characterization of the child's learning
strategy as "referential" or "expressive" (Nelson, 1973, p. 22).
Following Nelson (p. 17), we defined nominals as "general" if they
were the names for objects, animals, or abstractions, as opposed to
"specific," if they were used for just one exemplar of a category, as
with a proper name, or "blanket" to refer to the child's special
blanket. We made a new class, "social words," in which we
included specific nominals and Nelson's personal-social words
(assertions and social-expressive words like please or ouch) as well
as nursery routines, animal sounds, and greetings (but not func-
tional-relational words like gone, another, down, or most verbs, cf.
Gopnik, 1988). Following Nelson, we labeled children with more
than 50% general nominals in a language as referential in that
language and those with fewer than 50% as expressive.

A count was also made of the number and percent of verbs

Results

The 18 graphs of Figure 1 show the range of patterns of
bilingual vocabulary growth for the 18 longitudinal children as
measured by the CDIs between the ages of 8 and 30 months. The
bottom two lines are for single-language totals—how many words
in English and Spanish the child knew. The third line up is the
Total Conceptual Vocabulary (TCV), and the top is Total Vocabu-
lar y (TV), as described above. The area shading referenced to the
second y-axis represents the percent of time the child spent in the
two language environments: white for the Spanish environment,
gray for the English environment.

The following cues will help interpret the graphs. The
relation of the bottom two lines shows the relative dominance of
one language over the other. One can see directly which language
was stronger than the other and by how much. The distance of the
Figure 1A. Child 1 #6E. Language development in English, Spanish, and both languages together over time.

Figure 1B. Child 2. Language development in English, Spanish, and both languages together over time.

Figure 1C. Child 3 #6A. Language development in English, Spanish, and both languages together over time.
Figure 1D. Child 4 #72. Language development in English, Spanish, and both languages together over time.

Figure 1E. Child 5 #V1. Language development in English, Spanish, and both languages together over time.

Figure 1F. Child 6 #61. Language development in English, Spanish, and both languages together over time.
Figure 1G. Child 7 #61. Language development in English, Spanish, and both languages together over time.

Figure 1H. Child 8 #V7. Language development in English, Spanish, and both languages together over time.

Figure 1I. Child 9 #6E. Language development in English, Spanish, and both languages together over time.
Figure 1J. Child 10 #68. Language development in English, Spanish, and both languages together over time.

Figure 1K. Child 11 #69. Language development in English, Spanish, and both languages together over time.

Figure 1L. Child 12 #71. Language development in English, Spanish, and both languages together over time.
Figure 1M. Child 13 #V5. Language development in English, Spanish, and both languages together over time.

Figure 1N. Child 14 #6E). Language development in English, Spanish, and both languages together over time.

Figure 1O. Child 15 #67. Language development in English, Spanish, and both languages together over time.
Figure 1P. Child 16 #6C. Language development in English, Spanish, and both languages together over time.

Figure 1Q. Child 17 #V9. Language development in English, Spanish, and both languages together over time.

Figure 1R. Child 18 Comprehension #6G. Language development in English, Spanish, and both languages together over time.
TCV above the higher of the single languages tells how much the
minor language was contributing to the child’s conceptual vocabulary. If the child knew few words in the second language that she
or he did not know in the first, the TCV will be very nearly the same
as the line for the higher language. Similarly, the distance
between the TV and TCV reflects the child’s use of doublets: the
child with no doublet words will have identical values for TV and
TCV; children with more doublet words have a wider space
between the TV and TCV lines. (Note that the graphs tell who used
doublets and who did not, but the percentage of doublet knowledge
is underrepresented visually. The distance between the TV and
the TCV lines reflects the number of doublet pairs; i.e., approxi-
mately half the number of actual doublet words. The distance
between the lines has been further reduced by subtraction of the
phonetically similar words from the TV, as discussed above. Such
words averaged about 6% of TV for these children.) The left y-axis
of the graphs varies from 0–50 words in Figures 1D, 1F, and 1R to
0–800 words in Figure 1I.

We turn now to the results pertaining to the research ques-
tions raised in our introductory discussion.

1. How closely do the patterns of bilingual vocabulary growth
resemble patterns observed for monolingual children?

As can be seen in Figure 1, the sizes of the vocabularies of the
children at the different ages fell within the broad range described
by Fenson et al. (1991). Looking at their Total Vocabulary growth,
we see that most of these children started slowly and that somewhere
around the middle or end of the second year, the growth accelerated.
This pattern is most clearly seen in the 7 children with closely-spaced
data points between 12 and 24 months. Three younger children were
still in the earlier, slower phase at the end of the observation period.
Only 1 (CHI 13) of the other 15 children was actually observed to be
outside the range from Fenson et al. (1991) for producing 100 words
of Total Vocabulary by 27 months. Thus, the broad quantitative
expectations of vocabulary growth were met in this sample. (Pearson,
Fernández, & Oller, 1993, explored in greater detail how the bilinguals’
lexical growth related to norms for these ages.)
In terms of whether a lexical spurt occurred for these children, we had data from 12 children at the relevant time points. Three children (CHI 4, 6, and 18) had not yet produced 50 words and thus did not have enough words to have shown a spurt; one (CHI 10) had data points too far apart to allow a spurt to be observed. Two others (CHI 15 and 17) already had too many words in their dominant language at the time of the first observation, and so we could get information about the spurt or its absence from only one of their languages.

Following Goldfield and Reznick (1990, who were working from Gopnik and Meltzoff, 1987; also Bloom & Capatides, 1987), we accepted the definition of the spurt as 10 new words used twice in a 2½-week period. When adapted to our schedule of bimonthly observations and our method for counting words, that came to roughly 20 new words reported in a month. We looked, therefore, at whether any of the 12 children learned 20 words a month in each single language and in total language, especially when the children had fewer than 100 words. Using this definition, 3 of the 12 children showed a spurt in their major language: CHI 14 at 14 months, CHI 8 between 16 and 18 months, and CHI 9 after 16 months. A fourth child, CHI 13, spurted in English after 27 months, when his English vocabulary overtook his Spanish. Four children showed spurt-like growth if one considers not a single language, but growth in Total Vocabulary. CHI 1 and CHI 2 at 16 months, CHI 3 after 18 months, and CHI 5 between 23 and 27 months each added at least 20 words a month in both languages combined. This was true at these ages whether one counted Total Vocabulary or Total Conceptual Vocabulary (for all but one child, whose TCV growth was one word shy of the spurt definition).

No child showed a spurt in both languages at once, although one child, (CHI 9) did spurt in both, first in Spanish and then 2 months later in English. He jumped rather spectacularly from 33 to 403 words of Spanish between 16 and 20 months (or an average of 92 words a month). Then he "stalled" for about 2 months until he picked up in English, adding 90 English words and relatively few Spanish words in the next 2-month period. It is interesting to
note that the slope of the English spurt was not as steep as the earlier Spanish spurt, where he had increased both his Spanish singlets and TCV by an average of 100 words in four consecutive months. The relation of the TV and TCV lines tells us that the words he added during the English spurt between 22 and 24 months were mostly doublets, words whose concept was already known in the other language.

Finally, 4 children showed no spurt in either language. Three (CHI 7, 11, and 12) showed consistent, slower-paced growth in one language and even gradual loss in the other, the other child (CHI 16) showed little growth in either language. We have no data about whether the 2 children (CHI 15 and 17) who already had large vocabularies in their dominant language at our first observation experienced a spurt in that language. Despite having the 25 to 50 words usually associated with a spurt, their learning in their second language did not show any acceleration.

In this aspect of growth, then, bilingual children appear similar to monolinguals. Goldfield and Reznick (1990), for example, observed a spurt in only 13 of the 18 children in their sample. A similar percentage (8 of 12) in our study appeared to experience a lexical spurt. Whether the spurt happened in one language or both counted together appeared to depend on how rapidly the child was learning and how evenly the child's language learning time was distributed between the two languages.

2. How closely does growth in one language reflect growth in the other?

Few children in this study showed equal growth in both languages. Although all parents at the outset expressed a desire to provide the children with approximately equal input from both languages, one can see from the shading on the graphs a spectrum which went from balanced to unbalanced bilingual input and, hence, ability. In terms of producing a nearly equal number of words in each language, only 4 children (CHI 1, 2, 3, and 4), curiously with differing amounts of exposure to the languages, could be called balanced; another 9 showed growth in both languages but more in one than the other. For 6 (CHI 5, 6, 7, 8, 9, and
10) one language was more developed than the other, whereas 3 others (CHI 11, 12, and 13) experienced a change in which was the stronger language during the period of observation. Toward the other end of the spectrum, 4 children (CHI 14, 15, 16, and 17) looked largely monolingual, but nonetheless are included as bilingual as they had both English and Spanish words among their earliest vocabularies and continued growth, however small, in the minor language.

Looking at balance between the languages from the slopes of the lines rather than their relative heights, 3 children (CHI 1, 5, and 10) could be said to show parallel growth in the two languages. The 4 “near-monolinguals,” by contrast, exhibited strikingly non-parallel slopes in the two languages. For the other 11 children, growth appeared “complementary” in the two languages: changes in the slope of the line for one language may be said to be related to changes in the line for the other. For 8 (CHI 2, 3, 7, 8, 11, 12, 13, and 18) when the one language grew, the other slope flattened or even dipped; in 3 cases (CHI 4, 6, and 9) the total number of words declined temporarily as words were lost in one or both languages. In one of these cases (CHI 9), we saw one language standing still for several months while the other language appeared to “catch up”. The majority of the children, then, for whatever reasons, appeared to concentrate growth in one language at a time.

We originally wondered whether bilingual children would spurt simultaneously in both languages. After careful examination of these bilingual children, one can appreciate better how rare such an occurrence would be. To have a spurt in each language would require learning 40 new words a month, including at least 20 words a month in the second language as well. Only 3 children in this study (CHI 8, 9, and 17) added more than 40 new words a month, and as it happens, none of these three grew evenly in both languages, so only one language was credited with a spurt. Even the most lexically advanced child in the study, who experienced 4 months when he was adding 100 words a month to his CDI vocabulary, was not adding quite enough words in his minor language during that time to have it qualify as a spurt. A spurt
occurred in the second language a few months later, but only when
the first language was at a virtual standstill. Perhaps, with equal
support for both languages, a lexically advanced child learning 40
new words a month could have them evenly distributed between
the languages, but such an occurrence would require two rare
circumstances. Many more children will need to be investigated at
close intervals before we can draw any conclusions on this issue.
We can note simply that among these children it seemed to be the
exception to have a lexical spurt in both languages.

How much of the observed lexical growth was shared between
the two languages? Doublet vocabulary, that is, words known in
both languages, occurred in all the children, although individual
children varied in the number of equivalent terms they had. Once
the children had words in both languages, we can see that almost
all of the graphs in Figure 1 show some doublets, reflected in a
separation between the Total Vocabulary and Total Conceptual
Vocabulary (remembering that the distance on these graphs
underestimates visually the actual number of doublet words). The
consequences of the doublet vocabulary observed in these children
for theories of language acquisition are discussed by Pearson,
Fernández, & Oller (1994, in press). For the present discussion it
is important to note that a variety of patterns or strategies were
observed with respect to the establishment of doublet vocabulary.

Only 2 of the 18 longitudinal children showed a marked
preference for acquiring doublets, what one might call a "doublet
strategy". At 22 to 24 months, CHI 9 (discussed above with respect
to the lexical spurt) experienced a rise in Total Vocabulary and
English, while his Total Conceptual and Spanish measures re-
ained at the same level; that is, he was adding almost exclusively
the English equivalent for words he knew in Spanish. This child
was, as it happens, being reared in a situation that would encour-
age the learning of doublets: He was learning both languages in a
single environment, primarily from his mother. We know also
from the mother’s report that she was preparing him for a change
in his language environment; she was in effect teaching him to
switch his vocabulary to the other language. In addition to CHI 9,
CHI 18 between 15 and 16 months (in comprehension and at a much smaller scale) showed a comparable pattern of doublet growth (i.e., Total Vocabulary and vocabulary in English grew, while Total Conceptual Vocabulary remained relatively constant). None of the other 16 children showed a similar concentration on doublets, although everyone who was not actually losing words in a language was adding some doublets.

3. How differently can one characterize the growth in one language from growth in the other?

For a subset of the children, chosen to illustrate the range of different patterns observed, the type of words learned in each language is characterized in Table 1. As one can see, some of the children investigated were balanced between the languages in their linguistic development whereas others had a clearly dominant language.

Three basic patterns emerge from Table 1: similar proportions of referential words across languages for both “balanced” children (CHI 1 and 2) and those with a dominant language (CHI 17 and 6F); different patterns across languages for both balanced (CHI 5) and “dominant” children (CHI 11, 15, and V8); and a pattern developing from expressive to referential use of language observed in two languages simultaneously (CHI 9) and in only one language (CHI 12 and 14).

In general, as children acquired more words, the percentage of general nominals increased, so no vocabulary over 87 words was expressive. (Nelson’s, 1973, original labels applied to vocabularies only up to 50 words; in her study, an “object-naming” orientation at later points was determined with other measures.) Nonetheless, small vocabularies were not uniformly expressive nor larger ones uniformly referential. CHI 7 with only 15 words in English showed a referential pattern; similarly, CHI 5 at 23 months had relatively equal-sized vocabularies in each language, one referential and one expressive, and indeed, his Spanish vocabulary was still “expressive” at 87 words, 7 months later.

As expected, verb inventories seemed to develop when the child had a relatively large general inventory and when there was
<table>
<thead>
<tr>
<th>Child</th>
<th>Sex</th>
<th>Age in Months</th>
<th>Language</th>
<th>Number of Words</th>
<th>% General Nominals</th>
<th>Learning Strategy</th>
<th>% Social Words</th>
<th>Number of Verbs</th>
<th>% Verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>13</td>
<td>English</td>
<td>9</td>
<td>0</td>
<td>Expressive</td>
<td>100</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spanish</td>
<td>7</td>
<td>0</td>
<td>Expressive</td>
<td>100</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>Female</td>
<td>18</td>
<td>English</td>
<td>51</td>
<td>54</td>
<td>Referential</td>
<td>33</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spanish</td>
<td>56</td>
<td>55</td>
<td>Referential</td>
<td>29</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>Male</td>
<td>23</td>
<td>English</td>
<td>39</td>
<td>51</td>
<td>Referential</td>
<td>33</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spanish</td>
<td>32</td>
<td>28</td>
<td>Expressive</td>
<td>38</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>English</td>
<td>123</td>
<td>Referential</td>
<td>20</td>
<td>16</td>
<td>Referential</td>
<td>26</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spanish</td>
<td>87</td>
<td>39</td>
<td>Expressive</td>
<td>38</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>Female</td>
<td>25</td>
<td>English</td>
<td>15</td>
<td>60</td>
<td>Referential</td>
<td>40</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spanish</td>
<td>53</td>
<td>53</td>
<td>Referential</td>
<td>36</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>Male</td>
<td>16</td>
<td>English</td>
<td>27</td>
<td>15</td>
<td>Expressive</td>
<td>74</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spanish</td>
<td>33</td>
<td>18</td>
<td>Expressive</td>
<td>55</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>English</td>
<td>96</td>
<td>Referential</td>
<td>35</td>
<td>5</td>
<td>Referential</td>
<td>12</td>
<td>53</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spanish</td>
<td>401</td>
<td>56</td>
<td>Referential</td>
<td>12</td>
<td>53</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>English</td>
<td>Spanish</td>
<td></td>
<td>English</td>
<td>Spanish</td>
<td></td>
<td>English</td>
<td>Spanish</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>---------</td>
<td>-----</td>
<td>---------</td>
<td>---------</td>
<td>-----</td>
<td>---------</td>
<td>---------</td>
<td>-----</td>
</tr>
<tr>
<td>11</td>
<td>Female</td>
<td>26</td>
<td>19</td>
<td>Female</td>
<td>66</td>
<td>41</td>
<td>Female</td>
<td>39</td>
<td>32</td>
</tr>
<tr>
<td>12</td>
<td>Female</td>
<td>27</td>
<td>18</td>
<td>Female</td>
<td>22</td>
<td>32</td>
<td>Female</td>
<td>40</td>
<td>59</td>
</tr>
<tr>
<td>13</td>
<td>Female</td>
<td>28</td>
<td>26</td>
<td>Female</td>
<td>22</td>
<td>40</td>
<td>Female</td>
<td>42</td>
<td>32</td>
</tr>
<tr>
<td>14</td>
<td>Female</td>
<td>22</td>
<td>18</td>
<td>Female</td>
<td>33</td>
<td>40</td>
<td>Female</td>
<td>42</td>
<td>32</td>
</tr>
<tr>
<td>15</td>
<td>Male</td>
<td>21</td>
<td>26</td>
<td>Male</td>
<td>33</td>
<td>40</td>
<td>Male</td>
<td>42</td>
<td>32</td>
</tr>
<tr>
<td>16</td>
<td>Male</td>
<td>21</td>
<td>26</td>
<td>Male</td>
<td>33</td>
<td>40</td>
<td>Male</td>
<td>42</td>
<td>32</td>
</tr>
<tr>
<td>17</td>
<td>Male</td>
<td>21</td>
<td>26</td>
<td>Male</td>
<td>33</td>
<td>40</td>
<td>Male</td>
<td>42</td>
<td>32</td>
</tr>
<tr>
<td>18</td>
<td>Male</td>
<td>21</td>
<td>26</td>
<td>Male</td>
<td>33</td>
<td>40</td>
<td>Male</td>
<td>42</td>
<td>32</td>
</tr>
</tbody>
</table>
less emphasis on social words, onomatopoeia, and nursery routines. There were several exceptions, where children with relatively small vocabularies in a language had a relatively high percentage of verbs in that language (CHI 1 at 21 months and CHI 12 at 26 months in both languages, CHI 17 in Spanish, and CHI 15 in Spanish). These children knew more verbs per total words in that language than the 0% or 5% expected according to Fenson et al. (1991). For the most part, though, verb development in a particular language followed the path predicted by the child’s general development in that language. There are too few data, however, for statistical testing of this point.

Discussion

This study enlarges our perspective on bilingual development by moving the scope of the inquiry beyond the case study method. We viewed a restricted question, vocabulary development, from a consistent vantage point for a larger group of children than previously reported. Lexical development of typically developing monolingual children in the second year is known to be very diverse, and individual differences are great. When the differences in bilingual circumstances are added to the other social, cognitive, and affective factors which bear on language acquisition, the developmental picture becomes even more complicated. Almost all the logical possibilities regarding how the two languages could be related were exhibited in the growth patterns of these children. Nonetheless, several trends are apparent from the patterns of lexical development of these 20 children.

The rate and pace of development were similar to monolingual children’s. All but one of the 18 children with data at the relevant time points fell within the normal limits for acquiring early words. Such limits are still exploratory, but Rescorla (1989), for example, recommends a standard of fewer than 50 words produced or no two-word utterances at 24 months as a warning sign for expressive language delay. According to Fenson et al.’s
tables (1991), 5% of children produce fewer than 50 words at that age, so one child in 18 is as close as one could get to the 5% of monolingual children one might expect to show such a warning sign. In addition, phenomena of early lexical development in monolinguals, like the vocabulary spurt, occurred in this bilingual group.

Moreover, we observed a range of patterns in how closely growth in one language was related to growth in the other. These patterns no doubt reflect individual differences among children; however, more transparently than for monolingual children, the differences may reflect the children’s language environment—how the two languages were presented to them. Patterns of doublet learning, for example, differed according to the nature of the child’s language environment. The two children, CHI 9 and CHI 18, discussed above with respect to learning doublets, were both among the children who experienced a change of caretaker and hence language environment during the period of observation. Working backward from the environment, we proposed a way to predict which children would be most or least likely to establish doublet vocabulary: Children learning their two languages from two different individuals in two different settings should have the fewest doublets, followed by children learning their languages from two different individuals in the same setting; children learning the two languages from both monolingual and bilingual individuals in the same setting should have more doublets.

In a very general sense, most of the children with relative balance between their languages exhibited this pattern. CHI 2, 6, and 11, for example, who had different monolingual caretakers, had relatively separate vocabularies, with a good number of singlets in both languages. CHI 1, by contrast, who lived in a duplex with his English monolingual father, bilingual mother and sister, next door to his Spanish monolingual grandparents, had relatively equal numbers of singlets in each language, but also many doublets (at least at some observations). Like CHI 9 discussed above, CHI 1 was being raised in an environment that encouraged him to use two labels for the same objects or functions. A bilingual child from our
other research (#23, Pearson, Fernández, & Oller, 1994, in press) and another child reported on by Quay (1993b), also had bilingual and monolingual caretakers in one setting and a high number of doublets.

Without a principled basis for the expected number of doublets, we are unable to test the prediction statistically. In any case, the descriptive data suggest the importance of the language-learning environment in determining how independent one language will be from the other in the child's mind. (This aspect of bilingual learning for these children is explored further in Pearson, Fernández, Lewedag, & Oller, 1993; Pearson, Fernández, & Oller, 1994, in press).

What relationships between the two languages can be inferred from the separate graphs for the two languages? Parallel lines of growth might indicate an interrelationship between the two languages, or might simply reflect growth which is in fact parallel, but independent. A third possibility is that growth in both languages is tied to a common factor governing growth in both. Gopnik and Meltzoff (1987) suggested, for example, that the lexical spurt is related to changes that occur at about age 18 months in the way children categorize objects, and that success-failure words are acquired when children are successful at means-ends tasks. All other things being equal, one might expect a child who had achieved certain cognitive milestones to show reflections of that cognitive status in both languages.

Similarly, the occurrence of nonparallel growth lines might have contrasting explanations. Such asymmetry might signal the independence of one language from the other, or it might just reflect a difference in the opportunities the child has to bring the same language-learning capabilities and predispositions to the two languages. Finally, the complementary pattern of language growth observed for the majority of these children might be seen as an indication of reciprocity between growth in the two languages, or it might simply represent a limit on the children's processing capacity at any one time, causing them to focus on one language or the other, in turn. The quantitative data give no way to choose between these various alternatives.
Nonetheless, perhaps one can find some indication of the independence of one language from the other by examining the content of the child’s learning in each language. Even in cases where one language repertory is considerably smaller than the other, one can distinguish the child experiencing new growth even in the smaller language from the child whose one language appears to be completely subsumed in the other.

This question can perhaps be rephrased in terms of a contrast between first and second language acquisition. All of the children in this study met the technical definition of Bilingual First Language Acquisition (BF1A) stipulated by De Houwer (1990). All were exposed to two languages on a regular basis (i.e., they heard both languages every day) from birth on. Certainly they were all exposed to two languages well before the cut-off at age 3 years proposed by McLaughlin (1984) for simultaneous acquisition. Moreover, some of these infant bilinguals appeared to be experiencing first language acquisition in two languages, whereas others appeared to be mediating their learning of the second language through the first language.

One readily observed criterion for this distinction was the extent to which both languages contributed independently to the child’s inventory of lexicalized concepts. (The contribution of the minor language, beyond what was known in the other language, is seen on the graphs in Figure 1 as the distance between the middle two lines, the higher language and the Total Conceptual Vocabulary. Alternately, the independent contribution of the higher language can be estimated by mentally subtracting the number of doublet pairs, the distance between the two highest lines, from the second line. The remainder represents the singlets of that language.) Most of these children, once they had more than 10 words in all, had some singlets in both languages; that is, they knew some words in each language that they did not know in the other. Both languages were therefore contributing to the mental representation of the world that the child was in the process of using linguistic means to create. CHI 1, 2, 3, 5, 8, 10, and 11 all appeared to lexicalize new concepts in both languages. By contrast, CHI 14 and
17, and to a lesser extent CHI 15 and 16, learned almost no words in their second language that they did not already know in the first. On the basis not of the timing of their exposure to the second language but of its extent, these children may perhaps be characterized as infant second-language learners.

Whether BFLA applies for CHI 9 or CHI 13 is less clear. For each of them, after 24 months and after 27 months, respectively, English changed places with Spanish as the language where new lexical growth was taking place. Also not clearly in one category or the other, CHI 12 appeared to move from BFLA to monolingualism. In her case, English singlets were an indication not of new growth in English but of vocabulary loss in Spanish. All three children remind us that a static quantitative assessment will not suffice to make this distinction.

Two of the children proposed above as infant bilinguals experiencing second language learning also exhibit an interesting pattern on Table 1. CHI 15 and 17 had similar percentages of general nominals relative to social words in each of their two languages. Even the percentage of verbs, which for monolinguals has been linked to the number of total words (Marchman & Bates, 1991), was remarkably alike across languages, despite vastly different numbers of words in each language. The advanced cognitive development achieved by the child in learning the first language may be furthering learning in the second.

CHI 14 illustrates another possibility for the relation of the two languages: They appear to play different roles in her life. Like Bates’ daughter Julia (Bates et al., 1988, p. 257), CHI 14 seemed to use the second language mostly to engage in social interactions, as evidenced by the high percentage of her English lexicon derived from songs, games, and nursery routines. By contrast, the lower percentage of social words and the higher percentage of general nominals in her Spanish indicates a more cognitive orientation in that language.

Indeed when Nelson (1973) first proposed the idea of different strategies for learning to talk, she did not propose that one was superior to the other, just that they were different. Since then,
however, the referential style has been associated with faster vocabulary learning (Goldfield & Reznick, 1990; Horgan, 1981) and even grammar learning (Bates et al. 1988, Study 7). The young bilinguals in this study show some indications that a referential style represents more advanced language learning. Subject V8, for example, appeared to have, like CHI 14, a more referential style in her more developed language. CHI 12, despite a relatively high number of verbs, seemed quite delayed in both languages until 27 months, when she finally had enough general nominals to be classified as referential. CHI 11, too, was making more progress in her referential language. Similarly, CHI 5, who appeared to be a prototypical bilingual first language learner, showed more maturity in the verb lexicon in the language which also had more general nominals. The difference between his two languages may point to the language input as the source of the learning strategy; he was hearing English from his parents (both physicians) and Spanish from a nanny, who was perhaps more likely to use a socially-oriented language code with him. His example makes it clear that children can exhibit different learning styles in their two languages, even when these are relatively balanced.

Summary and Implications

This group of bilingual children exhibits patterns of lexical development similar in broad outline to the patterns observed in monolingual children. We observed that language development in the two languages for one individual can be quite divergent although parallels also often occur. Cognitive developments previously thought not to be language-specific are not always equally reflected in both languages. In addition, the range of different patterns demonstrated in this study is wider than those previously observed in case studies of bilingual children. We suspect that every possible relationship that can obtain between two languages in fact does, more even than can be demonstrated by these 20 children. A still larger number of children will have to be investigated for statistical trends to emerge.
These bilingual infants raise a new set of questions unique to the bilingual circumstance, but they are also a valuable source of information on general issues, providing a new perspective on questions like the origins of learning strategies—whether they are driven by the child’s preferences or by factors in the environment—and the relation between first and second language learning.

Revised version accepted 2 September 1994

Note

The CDI has shown high internal consistency, producing Cronbach’s alpha values of .95 for Infant Comprehension and .96 for Infant and Toddler Production. Test-retest reliability is also high, yielding Pearson coefficient values in the .8–.9 range for Infant Comprehension and Production and values exceeding .9 for Toddler Production (Fenson et al., 1991). In addition to demonstrating high reliability, the CDI has shown high concurrent and predictive validity (Dale, 1991; Dale et al., 1989). Dale tested concurrent validity correlations between CDI expressive vocabulary and performance on the Expressive One Word Picture Vocabulary Test (EOWPVT), the Index of Productive Syntax, and information obtained from language samples. Correlations to lexical and syntactic measures ranged between .68 and .78, p < .01: to the number of types (different words) in a 100-utterance sample, .74; to the IPSyn, .78; to the EOWPVT raw scores, .73, and to MLU, .68. Additionally, vocabulary scores from 228 children tested at two different times (Time 1—16–24 months, Time 2—22–30 months) were correlated at .71 (p < .0001), indicating relatively high stability, but allowing for differential growth over that period. Throughout the age range measured by the CDI/Toddler, correlations between successive ages are substantial and reasonably stable (Fenson et al., 1991).

Similar correlations for our sample of measures from 24-month laboratory samples to 2-year CDI production percentiles yielded the following values: to number of types in a 50-utterance sample, r(29) = .66, p < .001, and to the PPVT-R (Dunn & Dunn, 1981) at 30 months, r(20) = .77, p < .001.

References


toryl. San Diego: University of California, Center for Research in Language.


synonyms in the lexicons of bilingual infants: One language or two? Journal of Child Language.


