Applied Psycholinguistics 18 (1997), 41–58
Printed in the United States of America

The relation of input factors to lexical learning by bilingual infants

BARBARA Z. PEARSON, SYLVIA C. FERNÁNDEZ, VANESSA LEWEDEG, and D. KIMBROUGH OLLER
University of Miami

ADDRESS FOR CORRESPONDENCE
Barbara Z. Pearson, University of Miami Mailman Center, c/o English, Box 248145, Coral Gables, FL 33124

ABSTRACT
The bilingual child is seen as a unique source of information about the relation between input and intake. The strength of the association between language exposure estimates and vocabulary learning was examined for 25 simultaneous bilingual infants (ages 8 to 30 months) with differing patterns of exposure to the languages being learned. Using the MacArthur Communicative Development Inventories, standardized parent report forms in English and Spanish, the percentage of all words that were known in each language was calculated and then plotted against the estimates of language input (also in percentages). A significant correlation was found, $r(25) = .82, p < .001$. The correlation was also strong when examined point-by-point, even for children whose language environments changed by more than 20% between observations, although it was not reliable at lower levels of exposure to Spanish. Especially for children with less input in the minority language, the factors which appeared to affect the strength of the association between input and amount learned in a language are discussed.

In order for children to learn a language, they must be exposed to the language. In order to learn two languages, they must be exposed to both of them. Furthermore, the exposure must be direct, not indirect; watching people speak (or use sign language) is not enough to cause a person to learn to speak or sign. Rather, the learner needs to interact with speakers using the language (Ervin-Tripp, 1971; Griffith, 1985; Sachs & Johnson, 1976). These are the very basic conditions for learning language, and, indeed, no nonhandicapped human child for whom these minimal conditions are met will fail to learn to speak. These principles about language input are relatively uncontroversial, both in the linguistics literature and in the layperson's view of language acquisition. It is common sense to expect that the more a child interacts with speakers of a language, the more of that language the child will learn. But it is not obvious just how close the association between exposure and learning will be.

Environmental, or input, factors are generally thought to play an essential, but relatively small role in early language learning. In discussing language input, it is crucial to distinguish its role in learning a grammar versus

learning a lexicon. Even allowing for modifications of typical adult speech to promote children's learning, the so-called motherese (Newport, Gleitman, & Gleitman, 1977), input has been convincingly shown to be an inadequate stimulus for learning syntax (Chomsky, 1975; Pinker, 1994). Rather, a triggering mechanism is often proposed, whereby the individual child responds to a small amount of input provided at the proper time, just as kittens require exposure to light before 3 months of age in order to develop depth perception (Held, 1965; Scovel, 1988). In the principles and parameters framework for language acquisition (Chomsky, 1981; Hyams, 1986), for example, learners depend on input to learn how specific grammatical parameters are set in the language being learned. But as long as the amount of input is not reduced to zero, the effects of quantity are not considered relevant (Lenneberg, 1967; Pinker, 1994). Similarly, in examining the onset of basic language milestones, such as babbling with canonical syllables (Eilers et al., 1993), the first use of words (Fenson et al., 1991), and even the first word combinations (Hoff-Ginsberg, 1991), one sees, consistent with biological maturation explanations, remarkably narrow time frames across children and subcultures, despite vast differences in the amount of input provided.

Unlike explanations of syntax, parameter setting or other triggering-type mechanisms are rarely invoked as explanations of vocabulary learning. It is generally agreed that input enables word learning through some type of associative learning paradigm – in the past, stimulus–response, and now, perhaps, parallel processing models (Pinker, 1991). Even in the domain of vocabulary learning, though, input factors and growth are not always tightly linked. The "vocabulary spurt" is one widespread example where input quantity may be less a factor than the child's changing ability to make use of input. For most children, a bounded period of rapid vocabulary growth occurs after the first 25 to 50 words, but before the children turn their attention to syntax (Bloom & Capatides, 1987; Dromi, 1987; Reznick & Goldfield, 1992). The children's input conditions remain relatively constant, but changes in their cognitive state seem to heighten their ability to assimilate input (Gopnik & Meltzoff, 1987). So, the possibility that there is only a minor or indirect relation between input and vocabulary learning, especially in very young children, is quite plausible.

Research addressing the association of quantity of input with early vocabulary learning has a relatively short history. Huttenlocher, Haight, Bryk, Seltzer, and Lyons (1991) claimed to provide "the first direct evidence that amount of exposure is important to vocabulary growth" (p. 236). These researchers, studying 22 dyads of mothers and children between the ages of 14 and 26 months, found a substantial relationship between the amount that particular mothers spoke to their children and the variation in the children's vocabulary sizes. More recently, Hart and Risley (1995) presented a compelling amount of evidence suggesting that early vocabulary size is correlated with the number of words that children hear. Their ambitious longitudinal research recorded 42 children for an hour at a time each month for 2½ years. They transcribed not only the children's utterances, but also

those of the adults and other children interacting with them in their homes. Hart and Risley documented a strong, positive association between the size of the children's lexicon at each age and the number of words addressed to them by their caretakers.

Converging evidence of the effect of variation in input also comes from other avenues of research. Some support is found from cases of extreme deprivation, deaf children, and the occasional feral child (Tartter, 1986). In these circumstances, though, oral language input is so impoverished that the application of findings from these subjects to typically developing, hearing children has been difficult. Among typically developing populations, the effects of input quantity have sometimes been inferred from performance differences favoring first-born children: first-borns are considered to have more access to their parents' speech than later-borns (Hoff-Ginsberg, 1993). However, specific characterizations of the learning environments of first- versus later-borns are hard to develop. Other promising populations for direct tests of the association between input and learning are hearing children of deaf parents and young simultaneous bilinguals who, like first-borns, bring normal hearing and learning apparatus to the task. For these groups, normal language input (or, in the case of bilinguals, a specific language's input) can be present or absent in gradations across the whole spectrum of values between 0 and 100% of their waking hours. Both groups are considered to learn the full language (or languages) within normal ranges despite being exposed to the language(s) far fewer hours a day or week than children of hearing parents living in monolingual homes (Lenneberg, 1967; Penfield, 1967; Schiff, 1979; Western Pennsylvania School for the Deaf, 1992). Indeed, according to the widely cited claim from Penfield, bilingual children "learn two or three languages as easily as they learn one" (p. 193).

In a study of 52 hearing children of deaf parents, Schiff and Ventry (1976) looked at the relation between hours of exposure to hearing adults and children's language performance on a range of tests; they concluded that "amount of time spent with hearing adults did not appear to be an important variable" (p. 355). We know, though, that between-child effects on early language behaviors are extreme at these ages. Fenson (1991), for example, reported that the productive vocabulary of a group of 1,600 children had a standard deviation greater than or equal to the mean through 18 months of age. It is therefore possible that, with a group the size of Schiff and Ventry's, the between-child effect could be masking a potential environmental effect.

Children learning two languages provide a unique opportunity for researchers to hold the child-factor constant while testing variations in language exposure patterns. Certainly, there is no practical or ethical way to assign babies randomly to high- and low-input families. But the bilingual learner is her or his own "matched pair" (De Houwer, 1995), permitting the effect of environmental differences to be observed *within* children. In this way, individual differences in learning capacity or attention, for example, can be separated out from input effects.

Taking advantage of this ability to sort out child effects, 25 bilingual learners with varying amounts of exposure to Spanish and English were observed during the period of early language learning between 8 and 30 months. The particular study questions were as follows:

1. What is the strength of the relationship observed between language exposure estimates and vocabulary learning reported in that language?
2. What factors appear to affect the strength of that relationship?
3. Is the same relationship observed between input and vocabulary learning at all levels of input, or does there appear to be a threshold effect below which vocabulary learning does not take place, despite exposure to the language?

METHODS

Subjects

The data for this study came from 25 typically developing children (11 females, 14 males) being reared in English–Spanish bilingual homes in Miami, FL. Eighteen children were subjects in a 5-year longitudinal study on vocal development; in addition to the vocabulary measures described here, these children were tape-recorded in our lab on an approximately monthly schedule from 4 to 30 months and four times yearly from 30 to 60 months. Seven children from the university community participated in only the vocabulary segment of the study. All but two children came from middle-class homes. Four of the 25 children were born approximately five weeks premature with no other health problems. All vocabulary measures for the children were within the normal ranges for their age (Fenson et al., 1991; Pearson, Fernández, & Oller, 1993); in addition, a Bayley (1969) test at 18 months for 18 subjects averaged 113 (SD = 12.9). The data are essentially cross-sectional, but there are also longitudinal data from many of the subjects. The number of vocabulary observations varied from one to ten per child at approximately 2- to 4-month intervals between 8 and 30 months.

All of the subjects 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 eight households, both parents were bilingual. In 15 families, one parent and his or her extended family were native speakers of one language and bilingual to varying degrees, while the other parent was a speaker of the other language. In two homes, a nanny was the only source of the child's exposure to Spanish. For some children, the language environment was consistent over time, but in several 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 of the parents expressed a desire to provide an environment balanced equally between the languages (and the conditions of their households appeared to support that desire), in reality, only one child had equal exposure to both languages during the period of

observation. Parent estimates of language exposure, updated at regular intervals, averaged between 60 and 65% of one language and 35 and 40% of the other; four children had an average exposure less balanced than 75 to 25%. Eleven of the 18 children with more than one observation experienced a relatively consistent language environment throughout the data collection period; seven children experienced changes of 20% or more in the amount of time they were exposed to each language, four of whom experienced changes in which language predominated.

Materials

A standardized parent report instrument, the MacArthur Communicative Development Inventory (CDI), Toddler and Infant forms (1989), and its Spanish adaptations, the Toddler and Infant Inventario del Desarrollo de las Habilidades Comunicativas (Jackson-Maldonado & Bates, 1988), were used to assess the vocabularies of the subjects. The Infant English form contained 395 words frequently produced and understood by infants between 8 and 15 months of age, arranged in 22 semantic and grammatical categories. The Toddler form, used for children between 15 and 30 months of age, contained 679 words. The instructions on the Infant form asked parents to mark in one column the words their child comprehended only and, in a separate column, those words their child comprehended and spontaneously produced; on the Toddler form, parents were told to mark only the words that their child had produced. The vocabulary scores were the number of words marked by the parent, with one number for comprehension and another for production. All but seven observations reported here were for 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 listed 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).

Evidence of the CDI's reliability and validity is reported in Fenson et al. (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. In addition, two forms of this inventory focus on emerging behaviors at times when these behaviors are current (not retrospective) and limited in number. Its upper bound of 30 months reflects the fact that most typically developing children produce so many words by that age that parents can no longer keep close track of them. The CDI vocabulary list for a given child is not a true inventory, as the form does not exhaust the list of possible words that children might say. Rather, like vocabulary tests at later ages, it requires an extrapolation of the total vocabulary based on a controlled sample. In the research reported here, comparing one language to the other using the same measure was more important

than determining the precise number of words that the children knew. A careful study of the equivalence of the CDIs in the two languages showed that approximately 88% of the words reported for most of these children were words that were represented on both forms (Pearson, Fernández, & Oller, 1995).

Procedure

The children's parents filled out two CDIs for the child at a single age, one for each language. In some cases, one individual filled out both language forms for the child; in other cases, one parent did the inventory for one language and the other parent (or caretaker) filled it out for the other language. Parents also completed language background questionnaires. 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. Parents who completed more than one set of CDIs at different ages updated the language background questionnaire each time. An "observation," therefore, consisted of the pairing of a vocabulary measure and a language exposure estimate. There were 83 observations in all. The number of observation points per child depended on the length of time the family took part in the study, the child's age during the family's participation, and also on the parents' level of cooperation. Only three of the children were observed across the entire range from 8 to 30 months. Most (76) of the CDI measures reported here, for all but one of the children, were of productive, rather than receptive, vocabulary. The one exception did not produce any words in either language by 16 months when his family moved and he left the study. Since the principle is the same in comprehension as in production, we did not exclude from the analysis the seven comprehension measures we had from the children, although we did confirm that none of the results were substantially different when the comprehension scores were excluded.

As indicated in the CDI instructions, the parents were told to mark the 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 the child's production was different from the adult pronunciation. It should be noted that the information provided by the CDI was approximate in that the parents were not asked to specify the referent of a word. No claims are made in the use of the CDI that the children's words have identical meanings to those of the 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, two measures, English vocabulary and Spanish vocabulary, were taken directly from the respective CDIs. These were like monolingual measures except that each bilingual subject had two "monolingual" assessments. Wordforms used for more than one concept within a language (e.g., "ba" for "ball" and "ba" for "baby") were counted separately because they reflected two sound-

meaning pairings. When the same wordform was used across languages – even for the same referent (e.g., "wawa" for "agua" and also for "water") – it was counted in each language. Although its status as a separate lexical item is indeterminate from a theoretical point of view (Pearson et al., 1993), the production of such wordforms was considered to be an indication of the child's ability to function in each language environment. The two measures were summed, and then each language was represented as a percentage of the child's total number of words. The measures of language environment and the percentage of vocabulary in each language were correlated. (For a more detailed comparison of the children's cross-language vocabularies and their growth over time, see Pearson et al., 1993, 1995; Pearson & Fernández, 1994.)

RESULTS

Table 1 presents the data for each child at each observation, including the child's age in months, the number of total words (in both languages), the Spanish language environment, and Spanish vocabulary percentages. The children are arranged roughly in order of increasing exposure to Spanish, with the seven children with the larger changes in exposure percentages at the end (Subjects 19 through 25). Table 1 also notes the source of the exposure to Spanish. (The environments and vocabulary percentages for English are, of course, the complements of the figures given here.) The average number of total words for these children in these observations was 159, with a range from 1 to 778. Their level of lexical development was shown in previous work (Pearson et al., 1993) to be within the ranges of typically developing children given by Fenson et al. (1991).

Figure 1 shows the scatterplot of the percentage of Spanish vocabulary graphed by the percentage of Spanish environment. In general, the points cluster in the expected direction: that is, the more Spanish one hears, the more Spanish words one produces. As there are more points below the diagonal than above, there appears to be a tendency for a given amount of Spanish vocabulary to be associated with a greater percentage of language exposure than an equivalent amount of English; it appears to be harder for a child to learn Spanish than English, even within the Spanish-speaking community of Miami. This impression is confirmed in Table 1, where more of the points – for more children – show Spanish environment higher than Spanish vocabulary learned (counting cases with a disparity greater than three percentage points between the two measures). Individual data points on the graph can be located in the table, which is arranged as much as possible in order of increasing exposure to Spanish.

Correlations between environment and vocabulary percentage are shown in Table 2. (The seven observations with fewer than 5 words were omitted from this analysis. In theory, the principle should be the same: the greater the exposure, the more likely it is that words will be learned in that language. However, since the numbers are so small, the statistical effect of one word on the percentages becomes too great. Unlike the child with 100

Table 1. *Environment and vocabulary percentages*

Subject number	Source of Spanish	Age (mos.)	Number of total words	Spanish environment (%)	Spanish vocabulary (%)
1	mother	16	23	10	8
2	maid	21	173	20	17
3	father	26	330	30	21
		29	423	20	17
4	mother	14	3*	25	50
		16	22	25	78
		17	18	25	50
5	mother	18	73	40	11
		21	163	40	14
		24	264	40	20
6	father	10	9	50	30
		18	204	50	42
7	father & maid	11	0*	60	—
		11	112(C)	60	44
		16	60	60	65
		18	101	60	53
8	maid	23	65	60	45
		27	150	60	42
		30	200	60	41
9	father & maid	26	528	60	41
10	mother	18	20	70	81
		23	58	60	72
11	mother & father	9	2*	70	50
		12	15	70	67
		14	27	70	69
		16	51	70	53
		18	82	70	58
		20	121	70	64
		22	163	70	59
		24	214	70	51
		26	280	80	60
12	mother	30	698	70	62
13	mother	15	154(C)	70	75
		15	27	70	50
		18	127	70	68
		21	191	70	61
14	mother	26	537	70	63
15	mother & father	22	189	75	58
16	mother & father	15	10	75	60
		20	40	75	59
		25	65	75	78
17	mother & father	18	46	80	79
18	mother	12	18	85	75
		14	53	85	67
		16	122	85	79

Table 1. (*cont.*)

Subject number	Source of Spanish	Age (mos.)	Number of total words	Spanish environment (%)	Spanish vocabulary (%)
		18	207	85	86
		20	269	85	86
		22	334	85	89
		24	374	85	89
		26	419	85	89
		28	439	85	89
		30	460	75	93
19	mother	13	14	40	44
		16	14	40	56
		18	74	40	48
		21	136	60	58
20	mother & father	9	2(C)*	60	0*
		11	11(C)	55	62
		12	11(C)	85	58
		15	38(C)	60	74
		6	48(C)	55	62
21	mother & father	8	2*	50	0*
		10	13	60	64
		13	1*	80	0*
		15	15	50	65
		16	20	90	85
		19	32	80	64
22	mother & father	27	84	60	71
		30	150	40	41
23	mother	12	1*	75	100*
		16	51	75	55
		20	489	75	81
		22	455	40	81
		24	573	40	69
		27	775	50	56
24	mother	18	8	80	37
		20	40	80	71
		24	55	30	39
		26	87	30	30
25	mother & father	19	19	20	61
		23	54	25	40
		25	56	30	31
		27	76	40	13

Note: (C) in column 4 is comprehension vocabulary.
*Not included in correlation (fewer than five words).

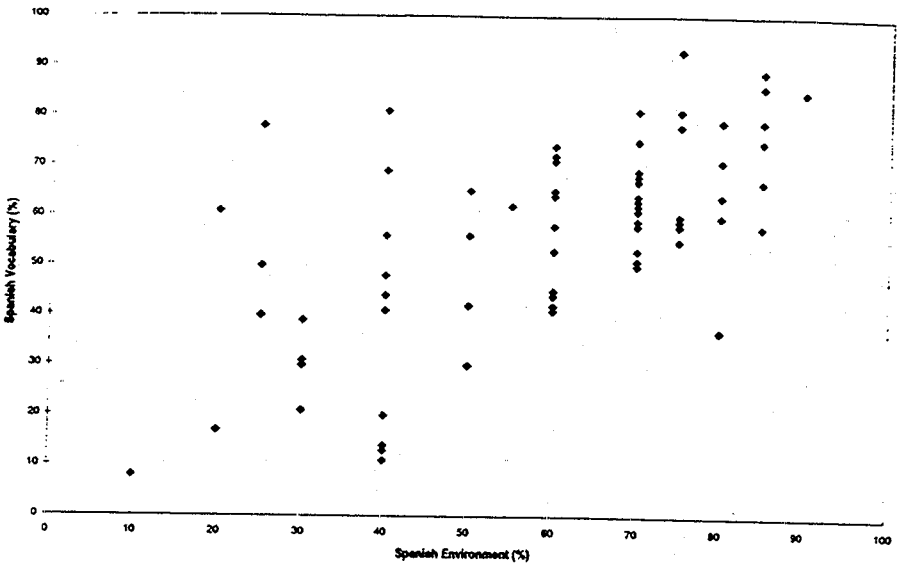

Figure 1. Vocabulary by environment (percentages).

Table 2. *Correlations between percentages of Spanish environment and Spanish vocabulary*

Group	Number of observations/ number of children	r	r^2	p of r^a
By child (average values)	/25	.82	.67	<.001
All observations (point-by-point)	76/25	.68	.46	<.001
Low Spanish (<50%)	22/10	.15	.02	>.10
Consistent environments (Subjects 1–18)	49/18	.77	.59	<.001
Changing environments (Subjects 19–25)	27/07	.49	.24	<.05

[a]Degrees of freedom for number of children, not observations.

words, where one less word in English changes the vocabulary percentage by one percentage point, the child who knows only 2 words goes from 0 to 50% on the basis of a single word.)

The first row of Table 2 gives the r, r^2, and p values for r for the group of 25 children, using a single value for percentage of Spanish vocabulary and Spanish exposure, which is the average of the longitudinal values for the child. Since our focus was on the association of the two figures – vocabulary and exposure – at any given observation point, irrespective of whether it was from a longitudinal or cross-sectional perspective, the second row shows the correlation for the whole group of 76 observations. For the

purposes of the point-by-point correlation, each observation point is considered as an independent event, granting that multiple observations of the same child violate the assumption of independence. Therefore, this and the following correlations are offered simply as corroboration of the preceding one. The *p* values were determined using the degrees of freedom from the number of children, not the number of observations. The third row shows the correlation for the 22 observations with Spanish exposure below 50%; the fourth and fifth rows focus on only those children with relatively consistent language environments (Subjects 1 through 18 in Table 1) and those with changes in the language environment greater than 20% between observations (Subjects 19 through 25).

It appears, therefore, that the relation between vocabulary learning and this relatively crude measure of language exposure is quite strong, accounting for two-thirds of the variance (67%) in percentage of vocabulary learned in Spanish when looked at child-by-child and for almost half of the variance (46%) in a point-by-point analysis. The relation is weaker for lesser amounts of Spanish exposure (in this case, the "minority language") and somewhat weaker, although still statistically significant, in environments that are not constant for the child. (The effects of such changes will be discussed later on a case-by-case basis.) Finally, there does not appear to be a threshold effect. There are nine children in our sample with exposure to one of their languages below 25% on at least one language questionnaire; an examination of their values on these two variables shows a comparable degree of association for most subjects, even with as low as 10% Spanish exposure (see Subjects 1, 2, 17, 18, and 21, as well as 20, 24, and 25, who are discussed later).

DISCUSSION

We have shown there to be a substantial relation between the quantity of input in a given language and the amount of vocabulary learning in that language during the second year of life. By comparing word learning across languages for a set of children learning two languages simultaneously, we have been able to show that the number of words learned in each language is, to a large extent, proportional to the amount of time spent with speakers of the language.

This would not be a surprising finding except in the context of statements by prominent scholars such as Penfield (1967) and Lenneberg (1967) to the effect that drastically reduced input has little or no effect on normal language learning. Given the strong tradition that, for so many years, has minimized the effect of quantity of input on learning, our finding of a strong relationship for early lexical learning is all the more noteworthy. Previous research with these same children (Pearson et al., 1993) gave indirect support to the suggestion that reduced input resulted in reduced productive vocabulary. In that study, the children's total productive vocabulary (whether calculated in terms of lexical types or lexicalized concepts) matched that of monolingual agemates almost exactly. Single-language in-

ventories for the children's dominant language were smaller than those for monolinguals, but given the wide variability among typically developing monolinguals, the bilingual means were not statistically smaller. This was not the case for the comprehension inventories; they appeared to be equivalent to those of the monolinguals in each language, but there were too few data to test that equivalence statistically. Nonetheless, the finding of this study and the earlier results, taken together, indicate that the bilingual's productive vocabulary inventory in each language at this age is some fraction of what he or she would have learned had more time been spent with speakers of one language and not the other. Of course, one cannot go so far as to conclude that the bilingual child's vocabulary learning in each language will always be a subset of the monolingual's, as "catch-up" points may come along later in development. Further research will have to address that question.

One assumption of this work is that percentage of exposure translates into amount of input. In fact, percentage of exposure could be rewritten in terms of the number of hours of exposure per day or week, with almost no change in the correlations. In our guidelines to parents for their estimates of exposure, we suggested that they work from an average of 12 waking hours per day. Thus, they counted the number of hours per weekday and per weekend day that the child normally spent with Spanish speakers and expressed it as a fraction of 84 hours (or whatever number represented the actual amount of time the child was awake).

Clearly, this figure does not translate into an absolute amount of input. There was some variability attributed to different sleep patterns among the children, but that variability was dwarfed by the much larger range in the number of utterances addressed to the children in their waking hours. We know from previously published and soon-to-be-published reports that the average number of utterances directed to a child in this age range, at least during the course of language collection studies, varied widely from child to child and, for the same child, in response to his or her stage of development. Hart and Risley (1995), for example, documented a rise in speech addressed to the child at around the time of the child's first words. From an exchange on the Info-Childes electronic bulletin board (15 June 1994), J. M. Siskind estimated that the average number of words that children in studies over the last 20 years had heard per hour was between 150 and 5,000. Rather than try to normalize observation times and volubility within that wide range at different ages (see also Hoff-Ginsberg, 1992), our study gave a *relative* measure. It is inevitable that some children experience a richer or a sparser linguistic environment in one of their languages compared with the other – a college-educated mother may be the source in one language and a poorly educated maid in the other. Nevertheless, the basis of the comparison was from one language to the total of the two languages for the same child. That is, whatever the actual number of words the bilingual children heard in Spanish or English, it was a subset of what they would have heard if 100% of their exposure to language had been in a single language.

It is an empirical question, though, as to what effect this reduction in input has on learning. There are several possible reasons why lowered input might not necessarily result in a direct, proportional reduction in the number of words learned. Perhaps children who are learning from reduced input are somehow more efficient at "intake," actually internalizing information, than children with more input; that is, they can derive equal benefit from less exposure. Or it could be that children receiving normal input cannot handle more than a portion of it, and the rest ends up as surplus, at least until they reach a later stage of development and are ready to assimilate more. A related possibility is that language learning does not take place continuously at optimum rates, but rather it happens in fits and starts, such that a deficit at one point in time could easily be made up for over the long run. If, indeed, children need only some fraction of the language input typically directed to them, their learning in each language would not necessarily be diminished by a reduction in the amount of input in each language, as long as the level remained above that fraction.

To date, no lower bound for sufficient input has been proposed, but a widely cited study of hearing children of deaf parents (Schiff, 1979) suggests that children spending 5 to 10 hours a week with hearing adults develop normally. Since all of these bilingual children were spending at least 8.4 hours (10% of 84 waking hours) in each language context, if they were learning two languages "as easily as one," we might expect "full" learning in each language. That would mean that each child would be learning however much she or he was capable of learning at that point in each language. In the terms of this study, full learning in two languages would translate into 50% vocabulary in each language, regardless of the time on-task. However, as can be seen from both the figure and the tables, this was clearly not the case. Fewer than a dozen of the 83 observations in Table 1 are close to 50%. That is, there did seem to be a direct effect of time spent in each language.

Of course, the correlation between input and intake is not perfect. Several factors could potentially diminish the strength of the relationship. First and foremost, one cannot ignore the potential inaccuracy of the parents' input estimates. Carefully planned work schedules involving monolingual caretakers mitigate against such error, but in the real world, people's lives do not always go according to plan. None of our families followed the "one-parent, one-language" strategy, which might have helped them quantify how much of each language was being spoken. Furthermore, bilingual speakers are not always aware of which language they are speaking (Goodz, 1989), and there is no control in a mixed environment over which language the child is paying attention to.

As seen in Table 2, the smaller correlation observed for exposure to the minority language (less than 50%) suggests that, at those levels, factors other than quantity determine whether input will be attended to. As Döpke (1992) and Lanza (1988) pointed out, the relative child-centeredness of the adult speakers of each language and the strategies adults employ to guide children's language choices are crucial for getting children to accept input

in both languages. The differing affective strength for the child of the language model in each language might also diminish the association of input quantity alone. For example, we observed two cases (Subjects 8 and 9) where the working mother spent less time with the child than the caretaker who spoke the other language, but the mother's language nonetheless seemed to exert a stronger influence on the language of the child's vocabulary. This was *not* the case, however, for Subjects 7 and 24, who also had working mothers, and so was not as great a factor in this study as it might have been. In the case of this group of children, one can see from the scatterplot in Figure 1 that the lowered correlation at the low end of the environment scale is the result of unexpectedly high amounts of vocabulary at low levels of exposure, as will be discussed later with respect to changes in a few children's level of exposure.

By making longitudinal observations whenever possible, we were able to highlight changes in environment as an important factor affecting the strength of the association between input estimates and vocabulary learned. Only one child (Subject 21) seemed to have suffered from the changes in language environment: she was as likely to lose words as to gain them between observations. But she was the product of a *very* inconsistent environment, living one month with grandparents and another with friends as her parents went in and out of jail and drug treatment programs. Language change was undoubtedly not the biggest obstacle to development she was facing. For the most part, though, the other children seemed able to respond to changes in their environment, but with a time delay. Therefore, one month's vocabulary reflected not the current language environment, but the previous one. In the case of Subject 20, changes in environment occurred at four observations in a row, and the vocabulary percentages followed the changes almost exactly, but a month later. Similarly, for Subject 23 the proportions of Spanish vocabulary at 22 months reflected the previous language environment. The change in the mother's input at 21 months (noted in the questionnaire at 22 months) and the fact of being cared for by the monolingual English grandparents for a month began to be reflected in the distribution of production vocabulary 2 months later, but it did not appear to have caught up with the change in environment until a full 5 months later. Translating from the scatterplot to Table 1, one can see that several of the points farthest from the diagonal in the top-left quadrant represent "out-of-phase" observations such as these.

Finally, parents sometimes reported changing the language environment in an effort to counteract a trend on the child's part toward one language or the other. Subject 11's parents were successful in encouraging more Spanish vocabulary at 26 months by increasing his time in a Spanish environment. By contrast, Subject 25's parents attempted to stem a decline in Spanish use, but their efforts seemed not to have been effective. Nor were Subject 18's parents able to encourage the use of more English. Both from the parents' report and the data in at 30 months, we observed the child's active rejection of her parents' efforts to introduce more English. The more En-

glish they spoke to her, the more she resisted. As it happened, the parents later learned that the child had a conductive hearing loss at that time, which seemed to interfere with her efforts in English but had a less noticeable effect on the more strongly established Spanish. Even after tubes were put in the child's ears, her hearing returned to normal, and her exposure to English was increased enormously, she still had more trouble than expected in switching to English. At the time of this writing, at age 4, her English is within the normal range, but is still markedly more restricted than her Spanish.

At the outset we had thought it possible that there would be a threshold of exposure below which vocabulary was less effectively learned. But for these children, that did not appear to be the case. Children with relatively consistent environments, with as little as 20% or less of their time devoted to one of their languages (Subjects 1, 2, 17, and 18), nevertheless still learned vocabulary items in proportion to the amount of exposure to that language. The lowered correlation at these very small amounts of exposure observed in this sample seems to derive mostly from the three children (Subjects 4, 23, and 25) who learned more than one would expect from their lowered exposure, not less. The only child who consistently learned much less than his exposure would predict (Subject 5) was working with relatively balanced input. So these data support the premise that, even at reduced levels of exposure to a language, children will still learn its vocabulary.

But what is the relationship between learning vocabulary and the more general sense of learning to use a language? At the earliest stages, the children's use of vocabulary via single words is the principal evidence of their active participation in the language community. Even at age 2, several of these bilingual children were still in the single-word stage in both languages; reported word use and language use therefore amounted to almost the same thing. Very soon thereafter, though, language use requires evidence of the development of syntax and language-specific phonology (Pearson, Navarro, & Gathercole, 1995). As children get older, vocabulary use alone becomes a smaller part of the constellation of skills required to credit them with "using a language"; vocabulary use will not as readily reflect language development more generally. Whether children can acquire a grammar and a sound system from low levels of exposure to a language cannot be inferred from vocabulary learning and will have to be specifically addressed in future studies.

The current analysis of these children's development provides the barest suggestion that children can learn the full language from less input, but with more difficulty. The fact is, the children hearing less than 20% of one of their languages were very reluctant to use that language in our lab, and they appeared in the play sessions to be "tuning it out." Recall that, with the exception of seven subjects, the children in this study were seen many times over a period of several years by the personnel of the Infant Vocalizations grant, who were responsible for eliciting tape-recorded speech samples in each language according to a schedule set out in the research design. The

research assistant and the parent spoke only one of the languages during each session, as stipulated. Not all of the children, however, cooperated, despite their being perfectly at ease with people and toys they had encountered all their lives. (Recording sessions began at 3 to 4 months of age.) That is, there were more than a few sessions where we were not able to collect a sample of 70 utterances in a language. Seven of the 18 children in particular posed a problem for our recording schedule. Six of them (Subjects 1, 17, 18, 21, 24, and 25) were reported to hear 20% or less of one of the languages, at least at the time of some of the vocabulary inventories. Only one child, who would not speak both languages for our recordings, was reported to hear a balanced amount of the languages (Subject 5, with 60% English, 40% Spanish). All of the other children generally obliged us by producing at least some utterances in the language designated for the session.

This pattern prompted the research staff to recommend that children whose exposure was less balanced than 75 : 25 not be considered for future bilingualism studies. This is not to say that some children with less exposure will not become bilingual or that all children with more exposure will, but that as a matter of practicality the elicitation of speech samples will be more efficient within those ranges. The implication for the current question is that we did not observe a lower limit on the percentage of exposure with respect to learning vocabulary, but we did see a limit below about 20% exposure in that the child appeared not to produce utterances in that language willingly or spontaneously. It will be an interesting assessment challenge, though well beyond the scope of this study, to probe the extent of the language-specific syntactic knowledge of children who respond appropriately to a language but do not produce any utterances in it (cf. Stromswold, 1994).

On a pragmatic note, judging from the experience of the families studied here, as well as those coming to our attention through the work of Döpke (1992) and Lanza (1988), it appears that, when bilingual children hear less of the minority language than they do of the majority language, parents may have to compensate for the difference with more active language teaching strategies than are normally associated with first language acquisition. Such teaching strategies are often developed naturally by some child-centered adults, but parents who are trying to provide bilingual experience for their very young children should be made aware of them and the importance of these strategies in achieving bilingual success, which has so often been taken for granted by scholars and laypeople alike.

ACKNOWLEDGMENTS

This research was supported in part by grant 5 R01 DC00484, NIH NIDCD and IR01 HD30762/01 NNIE-SBIR to D. K. Oller and R. Eilers. Portions of the article were presented in October 1994 at the Second Language Research Forum, Montreal, Quebec. We give special thanks to the Family Coordinator, Devorah Basinger; to the families who participated; and to our colleagues in the Bilingualism Study Group.

REFERENCES

Bayley, N. (1969). *Bayley Scales of Infant Development.* New York: New York Psychological Corporation.

Bloom, L., & Capatides, J. B. (1987). Expression of affect and the emergence of language. *Child Development, 58,* 1513–1522.

Chomsky, N. (1975). *Reflections on language.* New York: Pantheon.

(1981). *Lectures on government and binding.* Dordrecht: Foris.

De Houwer, A. (1995). Bilingual language acquisition. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language.* London: Basil Blackwell.

Döpke, S. (1992). *One parent, one language: An interactional approach.* Amsterdam: John Benjamins.

Dromi, E. (1987). *Early lexical development.* Cambridge: Cambridge University Press.

Eilers, R., Oller, D. K., Levine, S., Basinger, D., Lynch, M., & Urbano, R. (1993). Canonical babbling in infants. *Infant Behavior and Development, 16,* 297–315.

Ervin-Tripp, S. 1971. An overview of theories of grammatical development. In D. I. Slobin (Ed.), *The ontogenesis of grammar* (pp. 189–212). New York: Academic.

Fenson, L. (1991, June). CDI Basic Statistics: Final. Unpublished manuscript, San Diego State University, San Diego, California.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S., & Reilly, J. S. (1991). *Technical manual for the MacArthur Communicative Development Inventories.* San Diego, CA: San Diego State University.

Fernández, M. C., & Umbel, V. M. (1991). *Inventario del Desarrollo de las Habilidades Comunicativas, Adaptación Cubana* [Communicative Development Skills Inventory, Cuban Adaptation]. Miami, FL: University of Miami, Mailman Center for Child Development.

Goodz, N. (1989). Parental language mixing in bilingual families. *Infant Mental Health Journal, 10,* 25–44.

Gopnik, A., & Meltzoff, A. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development, 58,* 1523–1531.

Griffith, P. L. (1985). Mode-switching and mode-finding in a hearing child of deaf parents. *Sign Language Studies, 48,* 195–222.

Hart, B., & Risley, T. (1995). *Before Headstart: Differences in the family experiences of 1- and 2-year-old children.* Baltimore, MD: Paul H. Brookes.

Held, R. (1965). Plasticity in sensory-motor systems. *Scientific American, 213,* 84–94.

Hoff-Ginsberg, E. (1991, October). *Social class differences in maternal speech and their lack of consequences for children's early syntactic development.* Paper presented to the 16th Annual Boston University Conference on Child Language Development, Boston, MA.

(1992). How should frequency in input be measured? *First Language, 12,* 233–244.

(1993, March). *Object labels in early lexicons.* Paper presented to the Society for Research in Child Development, New Orleans, LA.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27,* 236–248.

Hyams, N. (1986). *Language acquisition and the theory of parameters.* Dordrecht: Reidel.

Jackson-Maldonado, D., & Bates, E. (1988). *Inventario del Desarrollo de las Habilidades Comunicativas* [Communicative Development Skills Inventory]. San Diego: University of California, Center for Research in Language.

Lanza, E. (1988). Language strategies in the home: Linguistic input and infant bilingualism. In A. Holmen, E. Hansen, J. Gimbel, & J. N. Jorgensen (Eds.), *Bilingualism and the individual* (pp. 69–84). Clevedon: Multilingual Matters.

Lenneberg, E. (1967). *Biological foundations of language.* New York: Wiley.

MacArthur Communicative Development Inventory. (1989). San Diego: University of California, Center for Research in Language.

Newport, E., Gleitman, H., & Gleitman, L. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children.* New York: Cambridge University Press.

Pearson, B. Z., & Fernández, S. C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language Learning, 44,* 617–653.

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning, 43,* 93–120.

(1995). Cross-language synonyms in the lexicons of bilingual infants: One language or two? *Journal of Child Language, 22,* 345–368.

Pearson, B. Z., Navarro, A. M., & Gathercole, V. M. (1995). Assessment of phonetic differentiation in bilingual learning infants, 18 to 30 months. In D. MacLaughlin & S. McEwen (Eds.), *Boston University Conference Proceedings* (pp. 427–438). Sommerville, MA: Cascadilla Press.

Penfield, W. 1967. The learning of languages. In J. Michel (Ed.), *Foreign language teaching: An anthology* (pp. 192–214). London: Macmillan.

Pinker, S. (1991). Rules of language. *Science, 253,* 530–535.

(1994). *The language instinct.* New York: William Morrow.

Reznick, J. S., & Goldfield, B. (1992). Rapid change in lexical development in comprehension and production. *Developmental Psychology, 28,* 406–413.

Sachs, J., & Johnson, M. (1976). Language development in a hearing child of deaf parents. In W. von Raffler-Engel & Y. Lebrun (Eds.), *Baby talk and infant speech.* Amsterdam: Swets & Zeitlinger.

Schiff, N. B. (1979). The influence of deviant maternal input on development of language during the preschool years. *Journal of Speech and Hearing Research, 22,* 581–603.

Schiff, N. B., & Ventry, I. M. (1976). Communication problems in hearing children of deaf parents. *Journal of Speech and Hearing Disorders, 41,* 348–358.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech.* New York: Newbury/Harper & Row.

Stromswold, K. (1994, January). *Language comprehension without production: Implications for theories of language acquisition.* Paper presented at the Boston University Conference on Language Development, Boston, MA.

Tartter, V. C. (1986). *Language processes.* New York: Holt, Rinehart & Winston.

Western Pennsylvania School for the Deaf. (1992). *Literacy initiatives for families of deaf parents with hearing children.* (ERIC Document #ED 353 391)