Finding Relations Between Input and Outcome in Language Acquisition

James L. Morgan
Brown University

J. N. Bohannon, R. J. Padgett, K. E. Nelson, and M. Mark (1996) raised numerous points of disagreement with the analyses of J. L. Morgan, K. Bonamo, and L. L. Travis (1995) concerning children's use of parental recasts, responses potentially providing negative evidence. Claiming to show that the bivariate time-series analyses used by Morgan et al. were inherently flawed, Bohannon et al. argued that training studies offer a preferable means for examining contributions of language input. In this article, multiple weaknesses of training studies are noted; at best, such studies may yield information on the sufficiency of aspects of language input but can yield none on their necessity. The failure of Bohannon et al.'s time-series analyses to distinguish among varying models of recast function is shown to be attributable to confounding of parameters and idiosyncratic assumptions adopted in generating simulated data from these models. Contrary to Bohannon et al.'s assertions, bivariate time-series analyses of observational data may provide invaluable tools for discerning signs and magnitudes of relations among variables in language development.

A child cries out at birth and would doubtless in any case after a time take to gurgling and babbling, but the particular language he learns is entirely a matter of environment. . . . The child learns to speak like the persons round him. (Bloomfield, 1933, p. 43)

Bloomfield's (1933) observations form part of the foundation of every theory of language acquisition. Because children learn only those languages to which they are exposed, defining the nature of this exposure is a necessary component of any account of how acquisition proceeds. But it may also be observed that, among organisms, only children succeed in learning language. Determining the nature of the perceptual, representational, computational, and linguistic capacities that enable children to perform this feat is thus also a necessary component of any theory of acquisition. It is clear that language develops as a result of interactions between the environment and structure in the language learner. A central problem for theories of language acquisition concerns how to apportion explanatory responsibility for the structure apparent in acquired grammars to these two sources.

Over the past several years, theorizing has perhaps tended to focus lopsidedly on defining linguistic constraints that might subserve acquisition. Several recent developments, however, augur increasing attention to the role of language input. These include the widespread availability of databases, including both parental and child speech (e.g., MacWhinney, 1991; also, Bloomfield, 1933, noted, "Almost nothing is known because observers report what the child says, but not what it has heard" [p. 512]), the advent of rich, multidimensional conceptions of information available in input (e.g., Morgan, Shi, & Allopenna, 1996), and the development of computational models for assessing possible contributions of such information to learning (e.g., Elman, 1990). In combination, these promise to foster the sort of precision, explicitness, and rigor with which hypotheses concerning language input deserve to be treated. Surprising and powerful results may emerge. For example, Shi (1994) has shown that, across a variety of languages, phonological and acoustic properties of infant-directed speech suffice to support inductive assignment of word tokens to basic grammatical categories with accuracies of 85%-90%. Elsewhere, of course, closer study may demonstrate that particular hypotheses concerning possible contributions of input are untenable.

Detailed analyses of input and models of theoretical learners may reveal whether in principle certain aspects of input are or are not necessary or sufficient to support particular aspects of acquisition. By themselves, however, these approaches leave unexamined a pair of critical issues, namely, whether children can and do exploit these aspects of input in acquiring language. One means of filling this gap is to pursue these issues experimentally, through training studies. In the following, however, I will argue that such studies suffer methodological limitations that severely constrain or cloud possible interpretations. In their place, I suggest that single-dyad correlational methods, such as those afforded by time-series analysis, may be equally informative and offer significant advantages of flexibility and economy. Developing, describing, and explicating such methods is thus an important priority for the field.

Consider, for example, how training studies might address a hypothesis such as the notion that parental recasts supply negative evidence that children use as a means of recovering from errors of overgeneralization. This hypothesis, which has important theoretical ramifications, has been championed by Bohannon and colleagues (Bohannon, MacWhinney, & Snow, 1990; Bohannon & Stanowicz, 1988) among others. One particularly interesting version of this hypothesis is that negative evidence supplied by recasts is necessary for acquisition. But one can see straightaway that training studies cannot be feasibly used for investigating this claim. What would be required would be to selectively deprive learners of any recasts. For obvious reasons,
this is impossible to do. A possible alternative is to study adults' or older children's learning of miniature artificial languages (see Morgan, Meier, & Newport, 1987, for a discussion of interpretive difficulties); another is to seek cultures in which recasts never appear in natural child-directed speech, as Gordon (1990) suggested. However, it is clear that training studies are confined to investigating whether aspects of experience are sufficient for promoting aspects of acquisition.

Training studies of language learning confront another serious difficulty: Acquisition phenomena often take months or even years to unfold. For example, my daughter began to over regularize past tense verbs shortly before her second birthday; as of this writing, she is 5½ years old and continues to overregularize. Given the time- and labor-intensive nature of training studies, it is prohibitively expensive to continue them for the period required to trace aspects of acquisition from beginning to end. Such protracted study, however, is precisely what is required for investigating hypotheses such as the one suggested earlier, where what is at issue is not whether recasts initially promote use of grammatical forms of constructions but rather whether recasts eventually extinguish use of ungrammatical forms.

Even short-term training studies run afield of methodological problems, a few of which are noted here. In general, training studies seek to manipulate frequencies of particular aspects of experience in order to measure effects of such manipulations on acquisition after some period of time. For training studies investigating language acquisition, such frequencies can be manipulated in one direction only, to increase those that occur naturally. To produce measurable effects over reasonable periods, frequencies may need to be greatly exaggerated. However, if, for example, very high levels of recasts are supplied for particular constructions that have been singled out for investigation (something that presumably never occurs in parental input), children may alter their productions of these constructions because of the demand characteristics involved rather than because of effects of recasts themselves. Second, because training studies of language learning investigate natural, ongoing processes, they are of necessity poorly controlled. Children who are exposed to varying frequencies of recasts within an experimental situation are contemporaneously exposed to varying frequencies of recasts outside that situation, and which of these is responsible for behavior observed at the end of a study may be unclear. In principle, frequencies of experiment-external recasts may be observed and used as covariates in analyses; in practice, this possibility has rarely been pursued. Finally, because trainers are necessarily aware of which treatment each child receives, the possibility that results may be tainted by experimenter bias cannot be discounted.

Careful design may minimize effects of these and other methodological problems, but instituting the required controls will further amplify the costs of conducting training studies. Moreover, the data acquired in any such study cannot be used to address alternative questions in language acquisition. In short, training studies require substantial, recurring investments and promise modest returns at best.

As an alternative, it may be more cost effective to pursue correlational studies of input and acquisition. Of course, the causal conclusions that can be drawn from any correlational technique are limited (finding associations qualitatively different from those required by a causal hypothesis may be grounds for discounting that hypothesis), but because no manipulation of any independent variable is required, such techniques can be adapted to a wide variety of existing databases to study a wide variety of questions. One correlational approach has been to measure input and grammatical growth over some lag for groups of children (e.g., Newport, Gleitman, & Gleitman, 1977). As discussed in the literature, this type of approach has certain weaknesses. For one thing, it tends to be Procrustean— as is well known, children develop language at different rates, and if growth is measured over a fixed period for all children, differences in developmental rate can either mask effects of input or appear as spurious effects of input.

More sophisticated approaches such as structural equation modeling might be contemplated, but these are not truly suitable for studying simpler issues of interest concerning whether the incidence of some particular aspect of parental input is predictive of aspects of children's later productions. Rather, given that both inputs and productions are distributed over time and are readily amenable to observation at regular intervals, bivariate time-series analysis provides an appropriate means for addressing such issues. Feldstein et al. (1993) described this procedure succinctly:

A theory of interactive behavior . . . must take into account how the person is affected by his or her own behavior (autocorrelation) as well as the partner's behavior (cross-correlation). But cross-correlations between the two streams of behavior may be an artifact of strong autocorrelation in each partner. Time-series regression analysis is a method that separately assesses and then removes (i.e., partials) the autocorrelation in each partner's contribution to the dyadic time-series. Then, by means of lag correlations (controlled for autocorrelation) it assesses whether each partner's stream of behavior is predictable from that of the other. (pp. 460–461)

Gottman and Ringland (1981) remarked

There may appear to be an indeterminacy in the model: The mother's behavior depends on her own past as well as on the baby's past, which in turn also depends on the mother's past. However, because of the assumption that the mother and baby are introducing independent innovations [i.e., that neither partner's behavior is completely determined by the other's], there are unique estimates for all parameters. (p. 401)

Several behavioral studies have used bivariate time-series analysis to examine microdevelopmental change (e.g., Feldstein et al., 1993; Field, Healy, Goldstein, & Guthertz, 1990; Field, Healy, & LeBlanc, 1989; Gottman & Ringland, 1981); these procedures can also be applied to series with larger temporal intervals, as is often the case in econometric studies.

Morgan, Bonamo, and Travis (1995; henceforth MBT) used bivariate time-series analyses to address part of the negative evidence issue, namely, whether children treat recasts as providing correction. If they do, one would predict that, for individual children, fluctuations in parental recasts of particular constructions at earlier times should be positively correlated with fluctuations in grammatical use of those constructions at later times. In other words, one would expect recasts to be positive leading indicators of grammaticality, much as building permits
are a positive leading indicator of housing starts. However, the results obtained by MBT failed to support this expectation. Although recasts were associated with grammaticality, the observed correlations were negative, at least for short lags, at which one may most plausibly expect to observe effects of episodic experience. Recasts are therefore negative leading indicators of grammaticality. This association is incompatible with causal explanation of recasts as corrections. Because this pattern was replicated numerous times by MBT, these results provide strikingly decisive evidence against the negative evidence hypothesis.

Bohannon, Padgett, Nelson, and Mark (1996) challenged MBT's conclusions on several grounds. The most cogent of these concerns whether the negative short-lag correlations observed by MBT might have been artifactual. In place of constructing an argument of logic showing that negative correlations necessarily arise from application of the analytic procedures used by MBT, Bohannon et al. resorted to attempting to demonstrate that analyses of simulated data sets, including those in which recasts have underlyingly positive effects on grammatical growth, yield negative correlations. Simulations can be a valuable approach to resolving such issues, provided they are pursued and reported with precision and explicitness. Unfortunately, Bohannon et al. failed to observe either of these principles; consequently, their critique is disappointingly uninformative.

Bohannon et al. (1996) appeared to base their simulations on a data set that bears little resemblance to any of those observed by MBT. They noted, "Out of roughly 1,200 article–noun errors in the Adam transcript ... target recasts [were] about 25% [of] maternal responses" (p. 552). However, MBT reported that "Adam's data included 2,253 grammatical and 2,077 ungrammatical noun uses" (Morgan et al., 1995, p. 183); the proportion of these that elicited minimal recasts (the response category found to be a leading indicator of later grammaticality) fluctuated and declined across time, averaging only 4% (parental recasting was not fully determined by children's errors). This distinctiveness is made clear by comparison of the results of MBT's and Bohannon et al.'s partial correlation results. Given that Adam's overall grammaticality was .52, on Bohannon et al.'s argument one would predict (according to the scattergram given in their Figure 1; see Bohannon et al., 1996) that a partial correlation of about +.40 should have been observed. The coefficient observed by MBT was −.55. All of the remaining grammaticality by partial correlation pairs observed by MBT would similarly be extreme outliers if plotted in Bohannon et al.'s scattergram. Apparently, the data observed by MBT might have been artifactual. In place of conclusions (as MBT did) that are spurious is vanishingly small.

Bivariate time-series analysis is by no means suitable for application to all child language transcript sets. The requirements of these procedures make clear that the sorts of databases that will be most useful are those containing large numbers of samples collected over very regular (and preferably quite short) intervals. These characteristics are particularly important for more informative cross-spectral analyses, which seek to uncover functional relations between inputs and productions. Until reliable speech-to-text devices become available, creating such databases will require very substantial efforts. Once such databases are in place, however, they can be interrogated through time-series analyses to address both questions of current import and questions that will assume importance in the future, as our understanding of both children's learning capacities and their linguistic environments continues to deepen.

References


depressed versus depressed dyads. Developmental Psychology, 26, 7–14.


Received December 1, 1995
Accepted December 1, 1995