

Cross-linguistic Study of Early Syntax

Diane Lillo-Martin

Abstract

The studies summarized here, as well as many others on-going in our laboratory, show that cross-linguistic studies of language acquisition can be a valuable source of information for linguistic theory. We welcome additional investigations of the principles and parameters of Universal Grammar from the cross-linguistic perspective.

Key-words: Language acquisition, syntax, grammar.

Resumo

Os estudos apresentados nesse trabalho, assim como outros em andamento em nosso laboratório, mostram que os estudos lingüísticos da aquisição da linguagem podem tornar-se uma fonte valiosa de informação para a teoria da lingüística. As investigações adicionais dos princípios e parâmetros da Gramática Universal sob a perspectiva translingüística contribuem para este estudo.

Palavras-chave: Linguagem aquisição, sintaxe, gramática.

INTRODUCTION

The theory of Universal Grammar is concerned with answering two basic questions (Chomsky 1986, p. 3):

- (i) What constitutes knowledge of language?
- (ii) How is knowledge of language acquired?

The approach to answering this question taken by generative linguistics has been to attempt to characterize those aspects of language which are constant across languages ('universals'), and those which vary. Concrete formalization of these properties provides the answer to (i). The answer to (ii) involves the postulation that the universal properties are innately given as Principles of Universal Grammar – part of the human genome. As such, the universal properties need not be 'learned', although they may require linguistic

input and/or biological maturation to be fully observed (cf. Crain 1991, Wexler 1999).

As for the properties which vary across languages, these are expected to fall into two groups. On one side are those properties which are simply arbitrary, and can only be obtained by exposure to them. For example, the fact that a particular organism is called 'tree' in English, but 'árvore' in Portuguese must be completely learned by exposure to these languages. On the other side are properties which fall under systematic, limited variation. For example, in some languages verbs systematically precede their complements, while in others the verbs follow their complements. These properties can be captured through the use of *parameters*, which provide alternatives from which languages may choose. Like the universals, the parameters are postulated to be innately given.

Parameters further contribute to language acquisition by connecting various effects under one umbrella. Languages in which verbs precede

Diane Lillo-Martin - professor at University of Connecticut and researcher at Haskins Laboratories

their objects typically display head – complement order in other categories as well, such as noun phrases and preposition phrases. Languages in which verbs follow their objects typically show complement – head order. Thus, determining the ordering of verbs and objects leads to knowledge about other categories, even without direct evidence of those categories. Theoretically, the connection between components of a single parameter may be quite indirect on the surface. This conception of parameters simplifies the task of language acquisition by reducing the choices which need to be made by children.

Under the view described here, the task of the language-learner becomes tractable. Rather than an indefinite search space in which the learner must somehow identify the target, the learner’s task is reduced to determining the proper settings of parameters, and learning language-particular facts such as vocabulary. In fact, it has been proposed that all variation, including parameter choices, is tied to particular lexical items (Borer 1983). Whether or not this can be maintained, it is clear that the ‘Principles and Parameters’ theory reduces the expectations on the learners as much as possible.

Why is such an extensive innate component required? The answer is clear when we consider the ‘Logical Problem of Language Acquisition.’ If the input provided to children were adequate for the target grammar to be reached, then perhaps a simple, general learning mechanism would suffice. However, the knowledge of an adult (i) includes the knowledge that certain strings of words are *not* part of the language. For example, consider the paradigm in (1)-(2).

- (1) a. I think Harry loves Sally.
 b. Who do you think Harry loves?
 c. Who do you think loves Sally?
- (2) a. I think that Harry loves Sally.
 b. Who do you think that Harry loves?
 c. *Who do you think that loves Sally?

In English, the complementizer *that* is generally optional in declaratives, as in (1)a and (2)a. When a *wh*-question is formed on the object,

the complementizer is still optional, as in (1)b and (2)b. However, when the *wh*-word targets the subject, the complementizer is not allowed – only the version with no complementizer is permitted, as the contrast between (1)c and (2)c shows.

This is not an isolated example of an obscure phenomenon. Native speakers of English clearly regard sentences like (2)c as unacceptable, regardless of their metalinguistic experiences.² And there are numerous other examples of such *constraints* operating in human language. Sequences that would be expected to be grammatical, on analogy with other grammatical strings, may be ruled out by constraints. How do speakers come to have this knowledge?

An obvious hypothesis is that speakers come to know that certain strings are ungrammatical because they are told so by their caregivers. This intuition is shared by many, particularly those who recall parental grammar correction from their youth. However, although parents may correct their children in some cases (I can clearly recall being corrected for using the wrong case on conjoined pronouns, e.g., ‘me and him are going’), they do not provide correction in cases like (2) – in fact, they don’t provide correction in such cases because children don’t even make such errors. Where parents do correct their children for ungrammaticality, they do so with such inconsistency and unclarity that children could not make use of this information (Marcus 1993, Morgan and Travis 1989). Further, cross-cultural differences in the type and quantity of corrective feedback to children make it clear that the answer to (ii) cannot depend on children receiving ‘negative evidence’ – i.e., explicit information that certain strings are ungrammatical.

Thus, the ‘Primary Linguistic Data’ – the input to which children are exposed – include examples of grammatically acceptable utterances,

²There are apparently dialects of English in which this kind of example is not rejected. However, other constraints are still clearly active in such dialects. The point is that all languages show evidence of constraints, leading to the problem of no negative evidence.

but not ungrammatical strings marked as such. Yet, the adult has knowledge of ungrammaticality. This state of affairs is schematized in Figure 1.

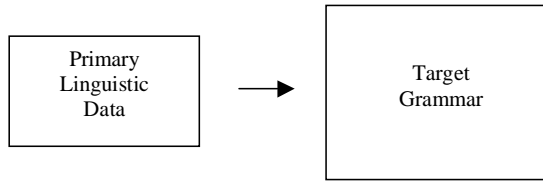


Figure 1. Primary Linguistic Data underdetermine the target grammar

The proposed solution to this problem adopted by generative linguists is that the innate Principles and Parameters of Universal Grammar provide the information necessary for the learner to achieve the target grammar on the basis of the primary linguistic data. On this view, the language acquisition scenario looks more like Figure 2.

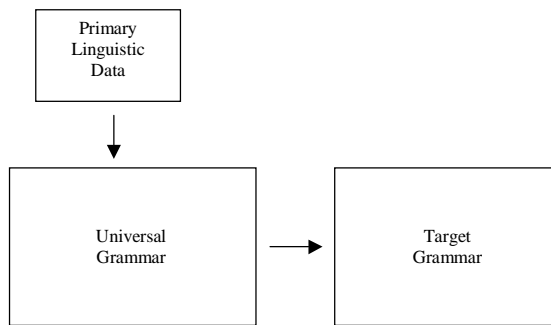


Figure 2. Universal Grammar solution to the underdetermination problem

What does this view predict as for the time course of the acquisition of various languages? Taken by itself, the generative theory is an idealization to instantaneous acquisition (Chomsky 1965). We know that the time course of acquisition is relatively short – even three-year-olds can be shown to exhibit knowledge of most aspects of the grammar they are acquiring; it is generally assumed that grammar acquisition is complete by age 5.³ Similarly, out of the many possible error patterns that children might evince, the attested patterns are very restricted. Thus, acquisition is relatively rapid and error-free –

³This means that the grammatical correction some adults recall their parents giving most likely arrived after grammar acquisition was complete, and the correction was aimed at achieving a more prestigious dialect.

though it is not instantaneous. Children do need to pick up on the existing differences between languages, and the time course of early acquisition can provide strong evidence about the organization of the grammar.

For example, children must determine the settings employed by their target grammar on various parameters. Suppose that children may temporarily entertain incorrect parameter settings. They may show evidence of this by producing utterances which reflect non-target settings. Importantly, such errors will not be completely random, but will reflect parameter settings made available by Universal Grammar and employed by non-target languages.

In a study which set the stage for research on the time course of parameter-setting, Hyams (1986) investigated the acquisition of null and overt subjects in English and Italian. In English, subjects of matrix clauses must be expressed overtly, as illustrated in (3). In Italian, however, the subject may be non-overt (null), as shown in (4). Information about the subject is carried in Italian by the presence of verbal morphology for person, gender, and number. Although verbal morphology marks third-person singular verbs in English, subjects must always be overt, as the unacceptability of (3) shows. This difference may be captured by a parameter (known originally as the Null Subject Parameter) by which languages choose to permit or to disallow matrix null subjects.

- (3) a. He speaks Italian.
- b. *Speaks Italian.

- (4) a. Lui (lei) mangia una mela.
 'He (she) eats an apple'
 - b. Mangia una mela
 'Eats an apple'
- (Hyams 1986, p. 31)

Hyams observed that young children acquiring both English and Italian frequently omit the matrix subject of their sentences. She proposed that for both groups of children, the initial setting of the Null Subject Parameter chosen is that which allows matrix null subjects, i.e., [+NS]. For English-acquiring children, positive evidence (in the form of overt expletives) would lead children to determine

that English takes the opposite setting [-NS]. On her account, differences between languages in the structure of modals also followed from different settings on the Null Subject Parameter – and as predicted, she observed that English-acquiring children only began to use modals when they ceased using null subjects.

Various aspects of Hyams' account of the early null subject phenomenon have since been amended, including both the theoretical analysis of languages which do and do not allow null subjects and the account of how children adjust their grammars. However, the approach taken by Hyams – finding concurrent acquisition of the various aspects of a parameter – has led the field.

CROSS-LINGUISTIC EARLY SYNTAX STUDY

In the Language Acquisition Laboratory at the University of Connecticut, we are studying the early syntactic acquisition of children exposed to American Sign Language (ASL), English, Japanese, Russian, and Spanish. This project, known as CLESS, constitutes an ambitious attempt to further the Principles and Parameters theory of Universal Grammar by collecting and analyzing comparable data across languages which share some features but differ on others. According to the theory of Universal Grammar outlined above, children should show evidence of adherence to universal principles from the earliest stages of acquisition. However, if children have not yet determined their target language settings on any parameters, they may show evidence of this through the use of non-target settings or through abstaining from producing elements governed by the parameters. We seek to examine children's acquisition of the various phenomena related to a single parameter in their target language, and to compare the acquisition of languages with similar or divergent parameter settings. In this way we test the Universal Grammar theory.

Another important aspect of the CLESS project, but one which will not be discussed at length here, is that we are including children for whom the timing of first linguistic input varies. For almost all children, linguistic input

(in at least one language) is received from birth. However, a large portion of Deaf children acquiring ASL do not receive input in ASL from birth. If the spoken language being used around them is inaccessible, the timing of linguistic input may be delayed by as much as several years.

To study the effects of different lengths of delay in linguistic exposure on the course of language acquisition, we are studying several groups of Deaf children. One group consists of Deaf children with Deaf, signing parents – these children receive input in ASL from birth. Other groups include children with hearing parents – one with relatively early input in ASL (by the age of two), and the other with relatively late input in ASL (after the age of five). We can examine the effects of linguistic delay by comparing these three groups with each other, and with the groups acquiring spoken languages (all with exposure from birth).

METHOD

The linguistic phenomena we are interested in studying generally become apparent in child language during the third year of life. At this age, it is extremely difficult to conduct traditional linguistic experiments such as elicited production, grammaticality judgment, or various comprehension tasks. On the other hand, for many phenomena of interest it is sufficient to see when they first appear in the child's production, what comes in at the same time, and what errors (if any) are produced. For other phenomena, it is of interest to see changes in production over a particular time period. Thus, a longitudinal sample of the child's language production in interaction with others is very useful for the present purposes. While such samples of spontaneous production are of limited use for some purposes (e.g., determining whether constraints on long-distance *wh*-movement are operative in the child's grammar), the phenomena we are currently studying occur with high enough frequency that this kind of information is ideal.

Our method is the same across the languages studied. We videotape children from



about the age of 1;6 (years; months) through 3;6 in order to capture the period before and after the explosion of syntactic development seen during the child's twos. The videotaping is generally conducted weekly, and sometimes semiweekly. The child is filmed while interacting with parents and/or research assistants. Because the children are visited frequently, they quickly get to know the researchers and interact with them quite naturally. The interactions are as natural as possible, with the exception that every effort is made to allow and encourage the child to talk. During these interactions, the children may look at books or play with toys – their own, or a set brought from the lab. One set of picnic toys, and another set of toy animals, has been used across all the language groups, to enhance comparability.

After filming, the tapes are transcribed in our lab, again using a consistent system across languages to the extent possible. All spoken languages are transcribed using CHAT format (MacWhinney 2000). In addition to standard tiers for the child and adult speakers, for phonetic specification of non-standard pronunciation, and for comments, child and adult utterances are paired with action and eyegaze tiers. This contextual information (about what the child is doing and where she or he is looking) provides crucial help for deciphering the probable intent of the child's utterance. A brief sample transcript of an interaction between DEB, the experimenter, and JOY, a child of 1;11, can be found in Figure 3.

```
*DEB: do you remember what's [: what is] inside the
apple?
*JOY: worm.
%pho: /woo6m/
%action: JOY turns around toward DEB
%eyegaze: JOY-off screen towards DEB (?)
*DEB: a worm.
@action: DEB reaches the apple and put her finger into
the hole of the apple
*JOY: a worm.
%pho: /6 wo6rm/
%action: DEB takes worm from JOY
*DEB: 0 [=! making playful sounds]
%action: DEB pops her finger out the hole in the apple
@Comment: DEB's finger forms the worm, it's a puppet
%eyegaze: JOY-the worm
```

Figure 3. Sample English transcript (DEB: experimenter; JOY: child, age 1;11)

For ASL, the transcription system is a bit different. A File Maker Pro database has been developed so that we can code all of the information used for spoken languages (and a report can be generated which employs the same

style as a CHAT file), as well as detailed information about the non-manual markers used – a crucial component for signed languages. In addition, the program Autolog (Pipeline Digital) is incorporated into the database. This program allows a field for the time-code 'address' to be included with every utterance record. When analyzing signed utterances, the researcher can use this address to cue the videotape to the spot on the tape when the utterance was made. Although there is sufficient information in the transcript so that we rarely need to go back to the videotape in analyzing spoken languages, for sign languages it is generally necessary to watch the child's utterance for a complete analysis. A sample screen shot of the database entry for one signed utterance made by the child JIL, age 3;3, is given in Figure 4.

Figure 4. Sample File Maker Pro record for one utterance (JIL: child, age 3;3)

RESULTS

The project described here is well underway, but still very much in progress. In this section, I will summarize two examples of the types of analyses we are conducting with these data, and their preliminary results.

Compounding and complex predicates. Snyder (1995, Snyder and Stromswold 1997, Snyder 2001) used the approach described here to investigate the occurrence of productive compounding and complex predicate types across languages. He found that there is a clear relationship between languages which permit productive noun-noun compounding (such as English) and those which



employ a variety of complex predicate structures, including verb particle constructions, resultatives, and dative constructions. He hypothesized that a common parameter underlies all of these types of complex word formation – and so, the timing of the acquisition of these constructions should be related. This prediction was confirmed. In particular, Snyder found that the age of acquisition of noun-noun compounding was a significant predictor of the age of acquisition of verb-particle constructions, datives, *put*-locatives, and causative/perceptual constructions for children acquiring English.

In the CLESS project, we have been continuing this investigation by looking at compounding and complex predicate constructions in Japanese. Japanese, like English, displays productive noun-noun compounding and various complex predicate constructions. Miyoshi (1999) and Sugisaki (2001) found the same relationship in the acquisition of Japanese that Snyder found in English between noun-noun compounding and causatives, double object datives, *put*-locatives, and resultatives.

Currently, we are extending Snyder's proposal by examining the acquisition of path and goal PPs in English and Spanish. Beck and Snyder (2001) found that the languages which permit compounding and complex predicates are also those which allow the combination of an activity verb and a goal PP to denote an accomplishment. Thus, this phenomenon would seem to be related to the complex word formation parameter uncovered earlier. A prediction of this hypothesis would be that the acquisition of goal phrases should be correlated with the acquisition of compounding in a language like English, which has the positive setting on the compounding parameter. As reported by Snyder et al. (2001), this prediction is confirmed. On the other hand, Spanish is a language which disallows complex predicate formation, does not have productive compounding, and does not interpret activity verbs with goal PPs as accomplishments. Apparently, children acquiring Spanish take some time then to determine its negative setting on the compounding parameter. We found that Spanish-speaking children fail to use goal PPs with *hasta* and *hacia* until well

past the age at which the corresponding constructions are acquired in English (Snyder and Lillo-Martin 2001).

Null arguments and verb agreement. Lillo-Martin (1991) found that there was a clear relationship between the production of verbal morphology and the use of null subjects and objects in the acquisition of ASL. In story-telling, four-year-old children who used uninflected verbs also failed to use null arguments, while older children were adult-like in their use of both verbal morphology and null arguments.

In the CLESS project, Lillo-Martin, Quadros & Mathur (1998) followed up on this study by investigating the acquisition of verbal morphology and null arguments in younger children acquiring ASL and Brazilian Sign Language. Unlike previous studies, this one found that young signing children employ verb agreement productively. As would be expected on an account which ties null arguments with verb agreement, the younger signing children used null arguments to accompany agreeing verbs.

It is well-known that Spanish is another language which has rich verb agreement and permits null subjects. It has also been claimed that children acquiring Spanish and Catalan (another null subject language) go through a very early stage in which they produce no subjects at all (Grinstead 1998). However, Puerto Rican Spanish (PRS) has lost most of the agreement morphology of Standard Spanish (SS), and it disprefers null subjects of most types. Since we have data on the acquisition of PRS as well as SS, Ticio (2001) has investigated this language change with respect to the optionality of subjects. She found, contrary to Grinstead's findings for SS, that children acquiring PRS show no stage completely lacking subjects.

ACKNOWLEDGEMENTS

The research reported here is supported by the National Institutes of Health, National Institute on Deafness and Communication Disorders, Grant #DC-00183; the National Science Foundation, Division of Behavioral and Cognitive Sciences,



Grant #BCS-0078788; and the University of Connecticut Research Foundation. The CLESS project is dependent on the collaboration of colleagues and students, as well as the participation of schools and families, whom I sincerely thank. I would especially like to acknowledge the contributions to the project of William Snyder.

REFERENCES

- BECK, Sigrid, and SNYDER, William. 2001. Complex Predicates and Goal PPs: Evidence for a Semantic Parameter. In *Proceedings of BUCLD 25*, eds. Anna H.-J. Do, Laura Domínguez and Aimee Johansen, 114-122. Somerville, MA: Cascadilla Press.
- BORER, Hagit. 1983. *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris.
- CHOMSKY, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- CHOMSKY, Noam. 1986. *Knowledge of Language*. New York: Praeger.
- CRAIN, Stephen. 1991. Language Acquisition in the Absence of Experience. *Behavioral and Brain Sciences* 14:597-650.
- GRINSTEAD, John. 1998. Subjects, Sentential Negation and Imperatives in Child Spanish and Catalan. Ph.D. Dissertation, UCLA.
- HYAMS, Nina. 1986. *Language Acquisition and the Theory of Parameters*. Dordrecht: Reidel.
- LILLO-MARTIN, Diane. 1991. *Universal Grammar and American Sign Language: Setting the Null Argument Parameters*. Dordrecht: Kluwer Academic Publishers.
- LILLO-MARTIN, Diane, QUADROS, Ronice Müller de, and MATHUR, Gaurav. 1998. Acquisition of Verb Agreement in ASL and LIBRAS: A Cross-Linguistic Study. *Poster presented at the Sixth International Conference on Theoretical Issues in Sign Language Research, Gallaudet University*.
- MACWHINNEY, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk. Third Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- MARCUS, Gary. 1993. Negative Evidence in Language Acquisition. *Cognition* 46:53-85.
- MIYOSHI, Nobuhiro. 1999. Compounds and Complex Predicates: Evidence from Japanese for a "Global" Parameter. In *Proceedings of BUCLD 23*, eds. Annabel Greenhill, Heather Littlefield and Cheryl Tano, 453-461. Somerville, MA: Cascadilla Press.
- MORGAN, J.L., and TRAVIS, L.L. 1989. Limits on Negative Information in Language Input. *Journal of Child Language* 16:531-552.
- SNYDER, William. 1995. *Language Acquisition and Language Variation: The Role of Morphology*. Ph.D. Dissertation, Massachusetts Institute of Technology.
- SNYDER, William, and STROMSWOLD, Karin. 1997. The Structure and Acquisition of English Dative Constructions. *Linguistic Inquiry* 28:281-317.
- SNYDER, William. 2001. On the Nature of Syntactic Variation: Evidence from Complex Predicates and Complex Word-Formation. *Language* 77:324-342.
- SNYDER, William, FELBER, Sarah, KANG, Bosook, and LILLO-MARTIN, Diane. 2001. Path Phrases and Compounds in the Acquisition of English. *To be presented at BUCLD 26*.
- SNYDER, William, and LILLO-MARTIN, Diane. 2001. On the Acquisition of Spanish Goal PPs. *To be presented at the 4th Conference on the Acquisition of Spanish and Portuguese as First and Second Languages, University of Illinois*.
- SUGISAKI, Koji, and ISOBE, Miwa. 2001. Resultatives Result from the Compounding Parameter: On the Acquisitional Correlation Between Resultatives and N-N Compounds in Japanese. In *Proceedings of WCCFL XIX*,



