SUSAN ERVIN-TRIPP

The Onset of Grammar

In the history of studies of child language, Ruth Hirsch Weir's book, 
*Language in the Crib*, will remain a classic. While many studies are su-
peredceded by later work with larger samples or more refined measuring instru-
ments, her book not only confronts important discourse problems for the
first time, but preserves in all its charm and vigor the freshness of the
primary data, the phonetic record of the child's monologues. These will
remain available for all future workers for testing new insights. Ruth Weir
was a pioneer of another new direction in child language, the live recording
of children's natural interaction. This work gives promise of further tribute
to her and of rich scientific rewards in the analysis of discourse structure
when the material is released. Until then, others will have to repeat her
labors.

The recording of children's texts has over the past fifteen years led us
far in the direction of new theories of grammar acquisition. It used to be
thought that children's first sentences were simply imitations of the surface
of adult sentences, and that the association of words with things, actions,
and concepts led to the eventual development of increasing skill in expressing
the child's ideas.

But studies of the tape-recorded texts of children's utterances like Ruth
Weir's have enabled researchers in the past decade to learn that children's
sentences often do not represent the order of the adult surface structure,
and that children's sentences have a systematically limited set of structures
which are simple and remarkably consistent. In adult speech we would call
such internal consistency evidence of a grammar.

From the more fragmentary diary case studies and from recent in depth
case studies from tapes of speakers of English (Bloom, Brown), Finnish
(Bowerman), Samoan (Kernan), Luo (Blount), Russian (Gvozdev), Serbo-
Croatian (Mikeš), Korean (Park), and other languages, we have learned some
ranges of variation and some common features to the onset of grammar.
The details of particular systems, like negation, plurality, or answering of
questions, have been explored in detail, and the evolution of the child’s
generalizations have been detailed.

The descriptive data collected in these studies, primarily based on text
collection in a fairly unstructured situation, was usually attacked as a problem
in grammar writing. The results of these analyses have shown some extensive
similarities between the types of sentences in the beginnings of multi-
morphemic output on both a structural and semantic basis. In my discussion
here, I shall be primarily concerned with the very early stages of grammar,
called by Brown (1973) Stage 1 and Stage 2, when there are around
1.35 morphemes per utterance. The major issues of language development
are already apparent at this age, but in highly reduced form.

In using as a basis for comparative studies the grammars written from
texts, we are plagued with the possibility that the grammars are somewhat
arbitrary. The strongest tool in linguistics — recourse to the informant’s
sense of grammatical acceptability — is not available with two-year-old
informants. For this reason we cannot know whether alternations like
book read and read book represent slips, equally plausible alternates, or
indicators of change in system. The informant’s judgments are in many cases
themselves a delicate performance, especially when dealing with alternates
with stylistic, sociolinguistic determinants or inherent variability. But they
may at least protect the grammar writer against the most outrageous ad hoc
inferences. We do not have that guide rail in child grammars. We invent units,
like the predicate or VP, but do not know if there is any such unit in the
child’s processing of sentences.

Further, child sentences are so regularly short that some linguists — e. g.
Bloom — have included brevity rules as transformations in the grammar.
There has been some disagreement about whether grammars should have
quantitative constraints, such as those on surface length, number of center
embeddings, and so on. There has been argument about how one can in
fact get closest to the competence of a child, as though there were some
fixed box with a grammar in it, distorted by the indirect access through
comprehension, imitation, and production. Most of us have regarded com-
prehension measures as somehow a tighter measure, less affected by variable
factors, probably because comprehension generally is ahead of production.
However, the fact that comprehension is extremely sensitive to semantic
contexts makes it probably less regularly governed by structural features
than production (Leontyev).

In accepting the traditional linguist’s task of writing grammars, workers
on child language have thus been saddled with special methodological
burdens. For example, in Ruth Weir’s book the phonemic categories judged
by standards of substitutability in context are different from the categories
judged in terms of substitution rules for adult forms (Ervin-Tripp 1964).
In a system undergoing rapid change, these inconsistencies are to be expected and should be predictable from an adequate account of the learning process.

This, then, is the goal of psycholinguists concerned with the ontogenesis of language — to go beyond the descriptive account of texts to examine the processes of language comprehension, and production, and how those processes change. To do this, we use data of many sorts.

In this paper I shall summarize certain of the descriptive generalizations we can obtain from the text studies of early sentences, and convert these into hypotheses about the process of development. There is considerable overlap between my summary and that of Dan Slobin elsewhere in this volume since we have been collaborating for some years in the same program of cross-language ontogenetic studies.

Prerequisites to language. Languages consist of patterned surface signals which are correlated, under particular circumstances, with extralinguistic events. In the adult language learner who hears talk about abstract ideas and about events at a distance in time and space, the meanings of utterances may appear to bear little relation to external events except when he goes shopping. Yet the learner must know the referent for learning of language to occur.

We shall consider here three categories of prerequisites to language learning. Two are those which the child brings to learning at the time we observe him — when sentence output is developing. These are strategies of perception and storage, and knowledge of concepts and relations. The third consists of what the environment must provide.

Cognitive relations refers to knowledge of the world which the child has gained from his interaction with his surroundings. It has been a striking finding of observers in a wide variety of languages that the structural meanings in the early sentences of children appear to be confined to a relatively restricted list. While samples in some languages have not included all of this list, the omissions may well reflect sampling deficiencies or cultural and situational biases which affect the child’s intentions while being taped. Here is a sample list:

**List A**
- Cigarette down.
- Give me candy.
- Candy mine.
- Hit you.
- Give me banana.
- You eat?
- I want water.

**List B**
- Girl rides.
- Bring candy.
- Baby’s eyes.
- Baby fell.
- Put down.
- Baby walks.
- Wants sleep.
Ball there.
Go home.
This visitor.

Keith there.
Go there.
Your baby.

It is not obvious that List A was collected in Kenya from Luo children (Blount) and List B in Samoa (Kernan).

a. Possession. In early sentences it is common to find utterances like candy mine, your baby, and baby’s eyes which are stable in the order of elements. The constituents are a possessor which usually is (+ animate), allowing for metaphorical extension to dolls and stuffed animals, and an object consisting of a part of the body, clothing, or a concrete object not inherently possessed. In the ontogeny of these representations, we may find Mommy when pointing to Mommy’s shoe, suggesting a loose associational relation.

If all that the possessive relation involved for a child was a loose association of possessor often seen with possessed, the order stability in the surface representations would be impossible, since there would be no designation of an asymmetry in the relation to identify an order.

b. Modality refers to the contrast between asking, demanding, and commenting. This difference may be signalled by gesture and paralinguistic features before identifiable words begin. Gruber claims that one child whom he studied intensively with particularly rich extralinguistic filmed information did not comment in the earlier stages of multimorphemic utterances, but always made demands, either that the listener look at or hand over.

Early sentences often contain explicit lexicon referring to questions (Where shoe?) or to demands (I want dolly), but where a single noun occurs it may not be obvious whether the utterance functions as a demand or identification or existence predication.

c. Vocatives occur early but are not strictly speaking relations. In many early texts we have examples of attention-getters, whether gestural by tugging on an arm, or vocal by calling a name. The name of the addressee or a vocative is often included in sentences having other functions and is typically not integrated grammatically.

d. Identification and Existence. Statements naming referents may be one-term nominatives, like dolly or they may contain more elements like See dolly, This dolly, That dolly. There is no evidence that these differences in surface representation represent semantic contrasts. This category is, of course, rich in cognitive antecedents since it requires isolation of objects and categories. Proper and common names appear undifferentiated at first.

Existence or “notice” utterances are situationally marked by a kind of discovery, but they do not appear to be marked by enough structural contrast to allow easy separation from demonstratives.
e. Nonexistence and negation. *No page, no rabbit, all-gone puzzle* are examples, and, of course, single negators in isolation precede. In the text of a deaf child’s early two-component sign utterances, Geballe found an absence of negators, but this may be idiosyncratic or due to small sample size. In Park’s text of a Korean child, there were a substantial number of negatives, both of the refusal and nonexistence type: *train there-is-not, no eat* (imp.).

f. Location. Locational notions are among the earliest acquisitions of the child, and in the course of language development locational questions and adverbs precede temporal expressions of analogous structural difficulty by years. The questioned location, as in *Where shoe*, to which the answer is pointing, is produced and understood from the beginning of questions. Early locational statements like *dolly car* or *go car* refer to places on or in which another thing or action is, or to destinations. Prolocatives like *here* or *there* appear in a variety of languages, but it is not evident that they are differentiated. Many forms used in English as verbs contain locational contrasts: *baby up* means to lift the baby up, *sweater off* to change the location of a sweater. *Fall-down* is an early English term requiring locational change in a specific direction. Given the presence of locations in identification and action sentences at Stage 1, it is not surprising that at the next stage when the surface structure is longer, locational information often is added to otherwise long strings, and that the development of prepositional differentiation of location occurs relatively early.

g. Recurrence of objects or events appears in *more apple, another toy, more throw*. These can be either descriptions or demands, and may occur without specification of object.

h. Attribution. Early sentences sometimes contain adjectives, as in *big truck* and *broken car*. Nominal attributives like *puzzle box* are, of course, homonymous with other constructions. The fact that English adjectives, which can appear either in predicate or attributive position, normally appear in child speech only in the latter position, suggests that in the input to children attributes are used primarily as identifiers rather than as new information.

i. Action-agent-object. Words describing actions occur early, though it is common to specify only one or two of the components of a three or four-part predicate like *put* or *give*, which have three nominal complements. Normally the agent is (+ animate) and the majority of objects, but not all, are inanimate. Bowerman found that subjects of action verbs might instead be vehicles. When expansion of noun phrases occurs, it is either in identification sentences or in objects.

j. States. Stative verbs like *want, like, see,* and *have* appear in the earliest sentences. It is not always obvious, of course, that they are semantically
clearly distinguished from active counterparts (see the book = look at the book). Nor do all children have such forms in the texts we have seen. Both the Luo and Samoan, but not the Finnish, texts had want, the most common of these.

K. Recipients. Datives or indirect objects can be found in early texts, though like attributives and states they may be sparse. Parisi and Antonucci (1973) have pointed out a parallel between the complement structure of verbs like give which require recipients, and put which requires a locative complement.

i. Conjunction. Pairing is usually of nominals without a marker. It is obvious that semantic development precedes its verbal expression and continues to grow independently, as manifested, for instance, in temporal categories. Slobin (1971) has asserted that these cognitive relations are not affected by the learning of language which merely provides representations for them. This may well be the case with universal categories, in which common conditions of life or at least of the interaction of child and milieu have been such as to guarantee development; in these cases it must be true that language is merely a manifestation. On the other hand, in cases where particular languages have categories that are idiosyncratic, the presence of the contrast might accelerate observation of a cognitive category or relation. An example is the study by Carroll and Casagrande of the shape categories codified by Navaho verb stems of handling and placement. These are high frequency forms in child speech and strikingly correlated with concrete categories of shape like long rigid, flat flexible, and so on. Navaho speaking children choose shape over color earlier than non-Nabaho speaking children from the same milieu — nursery school training in Boston had the same impact in creating form preference in English speaking children, so there are other ways besides language learning to create dimensional saliency.

The requirement that the child be able to identify “sames” in the sensory input simply is one other aspect of isolating recurrences with many transformations in the environment, but the fact that these events are auditory and involve changes related to features of the adjacent sounds, rather than spatial and light transformations, may alter what properties of perception and storage are required. Bever’s observation that long “jargon” utterances of a baby may be consistent situationally suggests that the first units may be fairly large. Since we are concerned here with grammatical acquisition, we will have to simply assume the child’s ability to process speech sounds in such a way as to identify and store recurrent patterns.

Environmental input. Since children learn to speak except under conditions of radical isolation, we assume that the input conditions necessary are relatively simple:

a. Orientation towards the signals. If acoustic input is relatively random in relation to events of importance to the child, as is the case if it consists
entirely of radio or TV broadcasts, the child may be unable to discover its structure. We have observed hearing children of deaf parents who had learned no speech from such input. The recurrence of signals at times of significance to the child, such as feeding, being changed, being held, being stimulated visually, may single out that channel for special attention.

b. Co-occurrence of speech with referential events. Let us suppose that parents held philosophical discourse in the presence of a baby, but never talked about the here and now. A child might in these conditions parrot gross features of their behavior, but the imitations would contain as many fortuitous as linguistically important features of speech. Thus a significant portion of speech must refer to concrete objects and events which are already singled out by the child.

Linguists argue that it is contrast of meaning which identifies which features of the sounds or of the grammatical patterns are significant; without such a criterion, the child would have no clue except recurrence to indicate which features must be stored as critical in absolute identification of items.

Is this to argue that nothing at all is learned from input devoid of meaning? Experiments with adults (see reference in Ervin-Tripp 1970) suggest that meaningless material, even when presented systematically so as to contrast acceptable and extraneous sounds, is destructive to language acquisition, possibly because the adults assign invented meanings to the forms. But these are adults. We do not know the extent to which hearing sentences without any referent can be in any way instructive about the structure of the sound system. The Clark and Hayes experiment with artificial sounds suggests that recurrent sequences may become salient and be stored even without the help of meaning, but their material lacked the phonetic diversity of actual speech since it was artificially produced.

It is certainly the case, however, that to learn grammatical features reference must be present for the new learner. For this reason, a considerable amount of the speech environment of children may be irrelevant to the learning of grammar, since the interpretation of none of the utterance is clear to him. Asher’s experiments with brief learning trials in which second language input co-occurred with actions by the learner to which the sentences or words referred suggest the dramatic effectiveness of making the co-occurrence of signal and referent optimal in time, in salience, and in sensory richness. The rapid fading of the acoustic signal may in fact make simultaneity important.

c. Recurrence of vocabulary in diverse environments; and moderate vocabulary diversity. In order to identify the more abstract units, starting with form classes, children must hear component units in the diverse positions possible. Not all lexicon, obviously, needs to meet this condi-
tion, and once classes are acquired a single encounter can mark new vocabulary appropriately. Conversely, varied entries in a fixed environment may be needed to teach formal features. This is less clear, since if entries share semantic features they may be spontaneously grouped as a class without this requirement.

In several observations of parental input in this country, it has been noted that some special conditions obtain which seem to contribute to the type of diversity experienced by the child. One is repetition, studied by Kobashigawa, who found that a relatively high percentage of input to children involved semantic continuity and considerable repetition of sentences with minor variations such as changes in intonation, word order, or optional deletions or contractions. In this condition, children would learn the semantic equivalence of formal alternatives readily.

A second input type is expansions, in which an interlocutor imitates the child by altering the child’s utterances in the direction of the adult model. As in the case of repetition, the child is thus given experience with adjacent formal alternatives while interpretation is held constant. In both cases, of course, there is no assurance that both parties agree in intention or interpretation. Where the child misunderstands and a repetition occurs, formal equivalence is still presented; where the adult supplies an inappropriate expansion, however, no clear grammatical facilitation could result. Indeed, in Casden’s experiment using expansions of all sentences, it was found that when expansion occurred for each utterance of the child, grammatical development was not accelerated as much as by talking to the child about semantically relevant matters. Possibly in the latter case there were fewer mismatches between the child’s intention or interpretation and the trainer’s, and the second condition does increase diversity of input.

Storage factors. If simultaneity of referential event and speech is a pre-condition of learning to understand speech, then it is necessary that one or both be stored for comparison. Young children’s own speech tends to refer to ongoing events like the naming of items or locations, or to immediately past events, as in dolly fall-down. When adults speak of distant events they may overtax the child’s ability to match meaning and utterance.

But even when utterances refer to the present, the rapid fading of the acoustic signal may require a short-term storage of acoustic information to allow work by the hearer in comparing the input with his referential knowledge. Given the small capacity in the immediate memory span of children, the selectivity of such storage is important in child language learning. Studies of imitation and of detection thresholds suggest the following generalizations about material with the highest probability of retention in short-term storage:

a. The most recent material. On the whole, spontaneous imitations draw selectively on the end of input material. Both trace and interference
theories of memory would predict the strongest control of the most recent input, given the fluctuating attention of spontaneous imitations.

b. Initial words. There may be some slight advantage of onset material over what is in the middle for unanalyzed new input, providing the child is already attending.

c. Vowel quality of stressed syllables.

d. Intonational contour, level of pitch, and other gross paralinguistic features.

e. Friction and nasality.

f. Order properties of sequential auditory material.

Since short-term storage is a prerequisite to building up long-term information about language, there will be prior acquisition of contrasts of the following types:

a. Prosodic contrasts and terminal juncture contrasts should be acquired early. In languages where modality contrasts can be signalled by prosodic cues, these occur before other signals like question words or inversions are used. In Japanese, Miyakara found that prosodic cues were used before the postpositional -ka.

b. Particles, enclitics, and items which are sentence final should be learned early before prefixes and other material earlier in utterances. In Japanese, yo and ne, which are clause final, are the first functors.

c. Suffixes should be learned more easily than prefixes and postpositions more easily than prepositions. The most striking support comes from theMikeš and Vlahović study in which bilinguals learned locatives in Hungarian before they used comparable prepositions making similar semantic contrasts in Serbo-Croatian. Slobin quotes Pačesová on Czech evidence that recency is stronger than stress as a determinant since initial stressed syllables are often omitted in Czech child speech.

A beautiful example of the intersection of two systems differing in the amount of suffixing is provided by Malmberg’s account of the acquisition of Swedish by a Finnish speaking child. The initial pattern by the child had of course Finnish noun stems and Finnish morphological material. The replacement was as follows:

1. Replacement of noun stem by Swedish noun.

2. The Swedish noun stem receives a Swedish suffix, followed still by the Finnish postposition as before.

3. The Finnish postposition is replaced by a Swedish preposition used following the noun and its suffix!

4. The preposition is correctly placed.

It can be seen here that the general order of putting in information in produced sentences is the hardest to change. The first shift is in the lexical material, as we would expect, rather than in inflections. The
Swedish suffix, rather than the preposition, is the next most noticeable element, and is readily produced because it can replace a Finnish unit in the same position. The preposition was the last feature, and curiously enough, though it was learned correctly, at first it was produced in the position of the semantically corresponding form in Finnish. The production of semantic information still retained a Finnish syntactic programming, though the morphological input was Swedish. In terms of the order of introduction of morphological elements, the order noun stem, then suffix, and, last, preposition is what we would expect on the basis of semantic and perceptual salience.

d. Sentence initial fixed-position forms should be learned more easily than material in the middle of utterances. In English, question words, the demonstratives, and prolocatives which often initiate descriptions are learned easily.

e. Morphemes which are syllabic should be learned more easily than consonantal morphemes. In the Mikeš and Vlahović study, the suffix -u rather than -t was used in both languages, then -ut!

f. Friction and nasal morphemes should be acquired sooner than stop or glide forms, unless there are articulatory problems in their production by the child.

g. Relative order of stem and affix should not be altered. There is no case in the child language data in which this has occurred.

h. Relative order of high frequency morpheme sequences should not be altered.

i. Where there is a dominant order for classes having structural meaning, when that construction is acquired it will reflect input order.

j. Unstressed syllables may be lost in storage of words. The Czech case cited earlier indicates, however, that recency may be a stronger factor than stress.

Long-term learning. Braine has proposed that recurrence prevents the decay of stored information, and that certain properties of the stored surface string when repeatedly processed through short-term storage will be preserved as the normal form of strings. Presumably, according to his model, this will happen whether or not the sentences have meaning, and his experimental work shows that adults can recognize novel but meaningless utterances conforming to a recurrent simple set of morphological rules.

Menyuk and Messer's child subjects reacted to violations of phonotactic regularities, so there are evidently some pattern constraints stored, as well as information about phonetic categories that allows recognition of accents. The extent to which this learning can occur apart from the acquisition of meaning is not known, nor have any of these studies been precise as to exactly which rules are learned first and in what form.
The fact that infants can make fine phonetic distinctions and recognize change between adjacent inputs tells us nothing about the issue of storage; long-term storage must require retention of partial information to make absolute judgments possible. How the critical information could be discerned in the absence of experience with meaning contrasts remains puzzling unless holistic, long-term auditory “Eidetic imagery” is possible.

Possible grammatical strategies. Grammar-recognition learning involves recognizing relations between order or functor patterns and properties of meaning. Braine’s work has focused on the recognition of order regularities and the conditions for such learning, without concern for the bearing that semantic correlates may have on facilitation of such learning. Semantic contrasts, of course, are only part of what is learned; it is a dramatic feature of both first and second-language learning that they go further than intelligibility requires in the direction of learning of formal correctness. In the case of children, meaning contrasts do seem important in relation to determinants of order of acquisition.

Random order. It used to be thought (see Burling, for example) that children’s first sentences were somewhat randomly ordered in terms of syntax. Suppose that the child discerns the relation between words and their meanings, but does not store any information about the orders of words, erasing what is in short-term memory about word order, but retaining order information as it affects morphemes within words. In the tests we have, we do not find random order. It seems unlikely that words are sufficiently distinct as units to children especially in the typical repetitive input to allow such a radical contrast between the type of information stored in the lexicon and in the rest of the storage device.

Preserve features on invariant positions. If a fixed morpheme like a question particle occurs sentence final, the modality contrast it signals may be recognized early. In the case of English question words, a group of words have nearly invariant position; in child speech, question words are always sentence initial, reflecting adult usage. We do not know whether relocating question words would interfere with their comprehension. In my work on answering questions (Ervin-Tripp 1970), the question word seemed rather like a dummy item with a fixed position for two-year-olds, in that children recognized that a question had been asked, but might answer as though a different question word had been used, the determinants of the choice depending on features of the rest of the sentence, like transitivity of the verb. The shared phonological features of question words in many languages may help in identifying them as such, at first, without any additional semantic features.

A storage device which marks certain lexical items as initial or final, or a sentence-producing device that is position specified has quite limited
utility. Speakers cannot in fact easily recall specific positions other than initial and final, so that only these positions could be indexed. The rest of the indexing would have to be relative rather than in terms of absolute position.

*Store relative order.* In languages where order carries semantic information, as in the contrast subject-object in English, the child must store instances like *pet the kitty* in order to allow him eventually to recognize that the item to be seen normally follows the action word, and to abstract the relation between action and object and match it with an order.

Highly frequent and adjacent orders in the surface structure appear to be learned quite easily, as we find in looking at possessor-possessed, agent-action, and demonstrative-nominal.

The fact that children produce sentences which deviate from adult order, like *all-gone toy*, of course, reflects overgeneralization, since the attribute-or quantifier-head sequence of the sentence is quite suitable to the majority of input cases.

*Store probabilities of order.* The surface order of units may not be consistent, either because it is affected by transform unfamiliar to the child (e.g. the English passive) or because the language does not in fact employ a consistent syntactic order in signalling a semantic contrast, as in the case of Finnish agent-action-object.

In such cases, the learner is evidently capable of storing a probability record. In the case of English, we find children who make occasional reversals of VO to OV, possibly reflecting passives or infinitive inputs, though passives are rare in input to most children in our samples according to Drach and Pfuderer. In Finnish, Bowerman reported a rough correspondence between the relative order for given structural types in the adult input and the child’s order.

**Structural Frequencies in Finnish Speech (Bowerman)**

<table>
<thead>
<tr>
<th></th>
<th>Seppo</th>
<th>Seppo’s Mother</th>
<th>Rina</th>
<th>Rina’s Mother</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject-verb</td>
<td>81</td>
<td>47</td>
<td>36</td>
<td>14</td>
</tr>
<tr>
<td>verb-subject</td>
<td>5</td>
<td>27</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>verb-object</td>
<td>24</td>
<td>16</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>object-verb</td>
<td>7</td>
<td>3</td>
<td>5</td>
<td>–</td>
</tr>
<tr>
<td>SVO</td>
<td>13</td>
<td>32</td>
<td>27</td>
<td>11</td>
</tr>
<tr>
<td>SOV</td>
<td>–</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>OVS</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>OSV</td>
<td>1</td>
<td>–</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>VSO</td>
<td>–</td>
<td>1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>VOS</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>noun-prolocative</td>
<td>13</td>
<td>14</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>prolocative-noun</td>
<td>19</td>
<td>41</td>
<td>97</td>
<td>31</td>
</tr>
</tbody>
</table>
This table makes clear that Rina's mother's most frequent construction was prolocative-noun, and this dominated for the child also. Seppo's mother's dominant type was SV, and this dominated for Seppo. The lower frequency of Seppo's SVO relative to his mother's is due to the fact he was at an earlier stage of language development than Rina, and used more two-word than three-word utterances. In each set of alternative orders, the dominant order is the same for adults and children but the alternative orders, if adults use them, appear in child speech. It would take fine grained contingency analysis of the texts to find if this order correspondence simply reflects some discourse echoing on a short-range basis, rather than any basic features of the child's grammar. In examining some order alternations in my own texts involving VO and OV changes, I found that what the adult said was often picked up by the child and therefore produced the appearance of alternation. In this way, I might be able to alter the frequency of passives in the speech of an interlocutor by raising my own frequency in a conversation.

Adult repetitions and expansions define equivalence. Where adjacent synonymous repetitions in the input are order preserving, whether the repetitions are of the speaker or are expansions of the child's speech, the child learns to preserve order in imitations and to retain order specifications in determining structural meanings. Where the parental repetitions or expansions freely transform order, as has been observed in Turkish and Finnish, children freely vary order in imitations. Whether there is a direct relation between this particular form of input and the child's processing of order information, we do not know. Such an experiment would require control over these varying types of input, including stability of order in all utterances of a given type addressed to the child.

What is the nature of the storage device? What does it store? One possibility is that it stores a word matrix. We know that speakers have remarkably good knowledge of word frequencies in the language (Howes). On this basis, maybe higher order matrices are stored. The strongest argument against such a device is its weakness. Such a device is too weak to generate novel sequences, like all-gone puzzle, which we know occur and have consistent grammatical properties.

If a word matrix existed, then the tallying of adjacent stressed words in texts of input to children would yield sentences like those found in children's speech. Here is a set of examples: Does go, Color my, Want hold, Come eat, Happened know, Going open, Diaper for, Animal this, Here big. To anyone familiar with children's sentences these sound strange.

It seems clear that many "prepackaged" sequences must be stored by speakers, appearing as phrases, idioms, cliches, whose production and comprehension can partially by-pass normal sentence processing routes.
Whatever storage device makes possible the retention of lexemes realized as words and phrases must also contain material which involves sequences heard more than once. In comprehension we have usually assumed that there is erasure or loss of surface information with retention only of semantic material. Sachs has shown that synonyms cannot be distinguished nor passives separated from actives after relatively small intervening material. If, however, it were the case that only semantic information is retained, language learning could not occur. There must be some storage of the phonological markers and semantic features inferred from the milieu of a new item for it to become part of the dictionary. It is not obvious that the device which accomplishes such storage has a strict word boundary and would refuse to store frequent sequences longer than words.

Semantic feature storage. In order to account for novel utterances, Schlesinger has proposed that the child stores the semantic features of utterances and abstracts structural rules from a series of instances. When he hears *Mommy's shoe, Daddy's shoe, Daddy's hat, Bobby's coat*, he already knows that Mommy, Daddy, and Bobby are people, that shoe, hat, and coat are moveable objects, and that in each case the specific object has a particular relation to the designated person. In this case the child would learn that Possessive relation → (+ Human) + (+ concrete object); the order would be a feature of the rule. According to Schlesinger's formulation, the Intention (possessive relation) is realized by possessor + possessed. If I interpret Schlesinger correctly, the components of structural rules are semantic features, and the categories are classes of items sharing these semantic features.

It is clearly the case that there are such semantic unities in the structural classes in early sentences. In one of Braine's lists of pivot classes we find a grouping of many words which can follow all: *broke, buttoned, clean, done, dressed, dry, fix, gone, messy, shut, through, we*. They all share a common aspectual feature and it is quite clear that the list is not interchangeable with a list of words following another pivot: *car, cereal, cookie, fish, high, hot, juice, read, sing, toast, walk*.

Melissa Bowerman noted that all of the items preceding action verbs in her corpus were either (+ animate) or (+ vehicle).

In my opinion the insight of Schlesinger's about the relatively large semantic homogeneity of early classes is important because none of the early grammars captured this property of child speech. McNeill's analysis of early form classes allowed him to move directly from formal classifications such as NVN to assumptions about semantic relations, though it is possible that NVN could represent location, not merely object, or a vocative rather than a subject.
There are several serious weaknesses in Schlesinger's account, as far as it goes, which also affect the attempt of Hebb, Lambert, and Tucker to subsume the ontogenesis of language under traditional acquisition of conceptual categories.

From early on, there are classes which are not semantically homogeneous, but which share considerable structural similarity. For example, we find both I want ball and Daddy throw ball. It is clear that throw and want have little if anything in common semantically, and Fillmore would argue that I and Daddy have quite a different role in the two sentences. Yet it would clearly be inefficient to write a grammar of English which did not subsume the two under the same major category, in view of their common destiny in tense change, and under a variety of transformations.

There are complex patterns such as the auxiliary system in English which are acquired in a relatively short period of time, obviously on the basis of transfer of patterns between items which do not share semantic features. I think it would be hard to find any semantic communality between can, will, and do which would be adequate to the rapid acquisition of the system by the two-year-olds in Boston and Berkeley. Yet Schlesinger's system, so far at least, would require that formal classes all be defined in semantic terms.

On the other hand, there are semantically related forms which do not have the same formal functions, such as eat, eating, and food. It is hardly a surprise that we need some formal marker systems such as affixes and functors to identify form classes, and that semantic unity is not enough. It is true that there are children like Park's, who in learning Korean occasionally employed words of appropriate semantic features in the wrong form-class position, such as throw for ball, and sit-down for bench. The fact that these occasionally occur makes us wonder why such substitutions are not much more common, especially in languages without class-marking affixes.

What could be the conditions for development of an abstract formal marker for lexical members of form classes? Morphemes and words and phrases which occur in the same environment, whether defined by surface forms or classes of forms, acquire a common formal marker. It is the case that those items occurring in the same environment will often have some semantic features in common, by virtue of the semantic co-occurrence constraints of the rest of the sentence.

The point where grammatical features must have been acquired is at the time when productivity occurs, and there is generalization from a semantically different item sharing distributional properties. It is the case, of course, that in most of the early grammars it would be possible to characterize the form classes by semantic features rather than a formal
term. Once formal features are acquired, many consequences follow — for example, the kind of associative behavior that Brown and Berko attribute to formal similarity of paradigmatic items.

McNeill has argued that some formal classes like noun and verb, are already present innately. Since languages differ in the formal classes required for their use, I should suppose that we must find some means by which classes are discovered. It is, of course, not necessary that the user of language in fact have all the formal classes linguists think will simplify the grammar. For example, many speakers may correctly employ verbal complements by learning them as phrases, just as we used to learn the preposition that followed French verbs, *reussir à*, and so on by rote. Unless there is a fairly rich set of syntactic consequences, if the only property shared by such covert classes is the particular structure or morpheme that complements them, there is, it seems to me, no need to abstract a grouping. I suspect that some adult speakers have never done so, and if asked to generate lists of "paradigmatic" items would follow semantic and larger form class constraints and not use the subclass list.

The richer the variety of structures in which a group of items occupies the same slot, the greater the likelihood that some formal marker will be generated, which in turn marks the lexicon and allow the structures to be organized as abstract rules, not low-level sequences.

**Why functors?** It is a striking fact of spoken languages that there always are classes of affixes or function words which are relatively small in number and high in frequency. Frequently these forms are not necessary for comprehension, as we find in listening to the telegraphic speech of children or of Japanese learning English. It seems to me that they have a function for comprehension, not for production, in marking form classes unambiguously, and it must be through these markers, whether or not the child at first reproduces them, that a considerable aid in generation of formal classes must come. The children in the Shipley, Smith, and Gleitman experiment who spoke in telegraphic utterances but understood normal better than telegraphic sentences were beginning to make use of this information.

It is relevant to this notion of the importance of surface markers to learning spoken languages that in pidgin languages derivational affixes may be created. In Neo-Melanesian, for example, the suffix *-fela* occurs for numerals, demonstratives, and one-syllable adjectives (Hall). Other affixes in Neo-Melanesian have more obvious semantic functions, e. g. *-[im]* to mark transitive verbs, as in [mi rid], [mi ridim b'uk], and [mi ridim], and the third person prefix, as in [ɛmi i-faytim ju'fela] 'he fights you-all', and [al i-krosim] 'they are angry at him'. My notion here is that oral languages must have surface bench markers to make complex sequences easy to process for the hearer.
The morphological examples bear on the low-level marking of form classes unambiguously, but it is also the case that functors may allow surface marking of more complex units. Bever, for example, as well as Shipley and Catlin working with children, have shown the importance of surface relative pronouns in facilitating comprehension of relative clauses, in contrast to deleted pronouns. But this does not concern us at the onset of grammar, and it is clear that while surface markers may facilitate this processing they are not absolutely necessary. The abstracting process which allows children to discover formal similarities is as early as the development of the auxiliary system in English.

Acquisition of affixes and function words. Like lexicon in particular sentences, functors differ in saliency. The factors which are likely to influence the frequency with which functors are kept in short-term storage and become candidates for the discovery of relations to extralinguistic conditions have been discussed earlier.

We can expect that if the semantic motivation for inflectional variation is not apparent (e.g. in the gender variation instance), the child may store the most frequent lexical form. He may learn *shoes* before *shoe*, but *dog* before *dogs*. In Finnish, verbs are first represented in the third singular form, and locatives in the allative or directional case, in Bowerman's sample. Park, on the other hand, found action verbs more often in the Korean imperative.

The first appearance of contrastive inflections seems to depend on the obviousness of the semantic contrast as well as the salience of the morpheme. It seems plausible that as early as a child could semantically contrast subject and object orders productively in English he might be able to use inflections for the same purpose, say in Caro (Burling); we do learn that Burling's child used inflections simultaneously with syntactic contrasts and that inflectional contrasts appear early in Finnish (Argoff).

The importance of the semantic function of the morpheme rather than its frequency is brought out by the example of *wa* and *ga* in Japanese. *Wa* occurs more frequently, yet it is *ga* that appears earlier in children's usage. Two differences in semantic function seem to account for the preference for *ga*. *Wa* can mark either the subject, the object, or any phrase to which attention is drawn. It is often translated as "as for X", so one can say "As for the table, we ate on it", or "as for John, he left", or "as for Mother's, it was lost". *Ga*, on the other hand, only can mark a subject, so that it differs in marking a class with a unique grammatical function. Further, rather than re-foregrounding understood, often antecedent information, *ga* typically identifies new information, and in that sense is likely to be more focal and more important. "As for the glass, it was the boy who broke it." In this sentence it is the boy, rather than the glass, which is new informa-
tion to the hearer, and we have glass-wa, boy-ga. The glass has been mentioned before, is presumably implicit, or was an implied contrast to some antecedent. The phrase could in fact be omitted if the presumption is clear. Thus both clarity of grammatical function, and novelty and focus would give preference to ga. With ga, there need be no ambiguity of subject and object, with only wa, there would be.

If the clarity of the correlation between the surface form and the meaning is obscured by the intersection of factors like gender which produce "noise", or by allomorphs related to phonological conditioning, then we would expect the most information-bearing contrast will be used and a simplification will occur selecting minimal features for the contrast. For example, in Serbo-Croatian the contrast between locative and directional suffixes is complicated by gender, so we have -i vs. -u as case contrasts, and -u vs. zero. We would expect that -i vs. zero might be selected to realize the contrast, since -u is ambiguous with respect to case.

The Language Acquisition System (LAS) must have at a minimum the following devices or processes, and we assume they apply to any language learning person.

a. Selective retention in short-term memory.

b. Phonological reorganization for retention in long-term memory.

c. Order retaining, which seems strong for acoustic material.

d. Feature template, which provides interpretations of structures according to the formal and semantic properties of sequences.

e. Formal feature generator, which identifies abstract classes and provides marking of the lexicon.

These devices need not necessarily be language specific, although at the moment it appears that order retention may be stronger for acoustic than for visual signal. Nor do I think these devices are age specific. The following factors will influence age changes in acquisition styles and rates:

a. Input conditions are quite different for adults than for children because external semantic reference of discourse will be reduced and the complexity of input linguistically will be greater with age.

b. Both immediate memory span and knowledge increase with age, making it easier to retain longer input and to discover meanings.

c. The system in greatest flux for a learner changes with age. By five, there is virtually no important phonological change occurring other than style and register enrichment. By ten, attention has shifted in aural communication almost wholly to semantic and lexical expansion. Attention to phonetic nuances other than to those carrying social meaning will have been sharply reduced, except in bilinguals. For these reasons, we can expect that the relative efficiency of acquisition devices dealing
The Onset of Grammar

with different aspects of language acquisition will change with age and with experience with language learning.

*University of California, Berkeley*

References

Antinucci, Francesco and Domenico Parisi

Argoff, H. D.

Asher, J. J.

Bever, T. G.

Bloom, L. M.

Blount, Ben G.
1969 "Acquisition of Language by Luo Children" (= *Working Paper No. 19* (Language Behavior Research Laboratory, University of California, Berkeley)).

Bowerman, M. F.

Braine, M. D. S.

Brown, Roger

Brown, R. W., and Jean Berko

Burling, Robbins

Carroll, J. B., and J. B. Casagrande
Cazden, Courtney B.

Drach, K.

Ervin-Tripp, S.

Fillmore, C. I.

Geballe, Carol
1969  "Early Sentences of a Deaf Sign Speaker". Unpublished term paper.

Gruber, Jeffrey S.

Gvozdev, A. N.

Hall, Robert A.

Hayes, John R., and Herbert H. Clark


Howes, D.

Kornan, K.

Kobashigawa, B.
1968  "Repetitions in a Mother's Speech to her Child". In Language, Society and the Child (= Working Paper No. 13) (Language Behavior Research Laboratory, University of Calif., Berkeley).

Leont'ev, A. A.
1969  "Inner Speech and the Processes of Grammatical Generation of Utterances". Soviet Psycho. 7: 3: 11–16.

Malmberg, Bertil
1945  "Et barn bytar sprak". Nordisk Tidskrift 21.
McNeill, David

Menyuk, P.

Messer, S.

Mikeš, M.

Mikeš, M., and P. Vlahović
1966 "Razvoj gramatičk kategorija u dočjem govoru". Prilozi, proučavanju jezika, II (Novi Sad, Yugoslavia).

Miyahara, Kazuko
1971 Language Development in a Japanese Child (Hakozaki, Japan).

Pačesová, J.
1968 The Development of Vocabulary in the Child (Brno: Universita J. E. Purkyně).

Park, Tchang-Zin

Pfuderer, Carol

Sachs, J. S.

Schlesinger, I. M.


Shipley, E., and J. C. Catlin

Slobin, D. I.