2 THE INITIAL EXTRACTION PROBLEM AND THE ONE UNIT STAGE

In the previous chapter I introduced the notion that the learning of "units," which may not necessarily correspond to the ultimate constituents of the adult's language, is the first step in a child's acquisition of the language, even though the recognition by the linguist of these units in the child's speech may present some methodological, but not insuperable, difficulties. What I wish to examine next are the circumstances surrounding the acquisition of these units. In particular, how do children get hold of these earliest units? Obviously they must extract them from the streams of speech that they hear around them. The problem of doing this is what I call "the initial extraction problem." In this chapter my focus will first be on what strategies children may use to attack this problem. Then I will look at what factors may produce the observed variation in individual children's solutions, that is, in the size and types of units they extract.

2.1. Extraction Of Early Units And The "Utterance Is A Word" Strategy

If language learners need to extract linguistic units from heard speech, what sort of knowledge and processing strategies must they bring to this task? In particular, do children have to know how to recognize what a "word" is in order to know how much to extract'? Or might they at first not be constrained to respect adult word boundaries'? In the light of the evidence already presented that early units may vary in size from words to whole phrases, I propose that children approach this task with the simplest strategy that can get the process started: They consider any utterance a potential lexical entry and copy and store it whole. The question what the child is to consider an "utterance" in an unknown language will be refined later in this section. It may not, as we shall see, always coincide with the adult's concept of an utterance .

Although the proposal that children may use whole utterances as their initial lexical entries has the advantage of allowing us simultaneously to account for both *one-word* and *formulaic* speech, the objection can be raised that many utterances are too long for small children to handle. Before proceeding with a discussion of other extraction strategies, let us first deal with this memory problem.

What would such an objection be really based on, besides our general awareness that early language productions tend to be severely limited in length? Is there evidence that young children's memories are not yet developed to the point where they can remember very many items (or very much of anything), whether nonlinguistic or linguistic'? Or is there evidence that children's memories are limited specifically with respect to linguistic material'?

There is some evidence that, as far as nonlinguistic material is concerned, younger children's memories are no more limited than those of older children. Thus Olson (1973, 151) reviews several experiments that involve children's recall of nonverbal materials (e.g., their memory for magazine pictures); when the children were tested in such a way that nonverbal cues could be employed in giving responses, younger children (around age 3) remembered as much as older children (of various ages). He concludes that younger children's apparent memory limitations are due to an inability to encode the responses verbally rather than to an inability to remember numbers of items.

Snow (1981b) also reviews evidence that young children are perfectly capable of remembering complex events for long intervals. For instance, DeLoache and Brown (1979) found that 18-month-olds could remember the locations of toys for several hours. Even more compelling is Snow's diary evidence: Her 2-year-old son was observed to remember very specific (but mostly trivial) incidents over periods of time ranging from three days to two months.

If children's memories for nonlinguistic material are not limited, what about linguistic material? It has in fact been observed that even young children can be very good at remembering and reproducing fairly long chunks of speech. Thus, Bloom notes,

the children's ability to remember complex linguistic material was often impressive as evidenced by the production of stereotype model sentences that had no analogue in the child's grammar; for example, "who has that book?, Kathryn has that book" at Kathryn I [1;9]. An even more striking example was Eric's ability to recite" accurately from memory long passages of text in his favorite story books. turning the pages at

the appropriate juncture, when he was two and one-half years old. But he was unable to answer specific questions about the text, and certain phrase structures he recited sentence adverbials for example - did not appear in his spontaneous utterances.   [1970. 168-9]

The evidence therefore shows that some children can in fact remember whole utterances, some of which can be quite long. It does not seem unreasonable, then, to attribute to all children the strategy of extracting at least selected whole utterances.

In order to make more concrete the proposed extraction process, I will present it as a series of heuristics for processing linguistic input. These heuristics are modeled after Slobin's Operating Principles (1973) for acquisition of syntax, but they apply to a much earlier stage of language acquisition. The two most basic heuristics are

EXTRACT1. Extract whole utterances from the speech stream surrounding you and remember them together with salient features of the situational context. EXTRACT2. Store the whole unit as a single entry in your lexicon.

Determining what is an "utterance" is influenced by saliency properties of the input speech and by early limitations on the amount of information the child is able to pay attention to at once. These constraints make it necessary for the child to enlist further heuristics to focus attention on selected aspects of input speech, in particular to delimit stretches of speech and to select which of these are to be remembered.

One heuristic that limits the choice of utterances is based on saliency of meaning:

EX:MEANING. Extract and remember sound sequences that have a clear connection to a clear context.

The problem of delimiting utterances (i.e., chunks to be extracted) is dealt with by a set of heuristics that make use of salient properties of the input speech. Probably the most salient property is the distinction between silence and speech. The first delimiting heuristic I propose is

EX:SILENCE. An utterance is bounded by silence.

Children have been observed to extract, presumably on the basis of this heuristic, single words (when they are bounded by silence), whole phrases (bounded by silence), and even two run-on phrases (bounded by silence but not sufficiently separated by silence). These three cases are respectively illustrated by the following three examples from the literature.

1. Nelson (1973, 104):

    Mother: Jane. Here's a bottle. Where's the bottle? Here's a bottle.
    Jane: Wah-wah.
    Mother: Bottle.
    Jane: Bah-bah.

2. Peters (field notes):

    Minh's mother frequently exclaimed "Look at that!" to him and [dukə'dæt] was one of Minh's first utterances that I could associate with any certainty with a known target.

3. Clark (1977, 350):

    Adam: Squeeze.
    Adult: Squeeze? You squeeze.
    Adam: Squeeze you squeeze.

Another salient phonological property is intonation contour. It is used by the following heuristic:

EX:SUPRASEG. An utterance is a suprasegmentally delimited stretch of speech.

In her study of the prosodic characteristics of American mothers' speech to 2-year-olds, Garnica found that more sentence-final pitch contours were rising than in these same mothers' speech to 5-year-olds or adults. She hypothesizes that these rising pitch terminals "may be used to cue the child to the location of sentence boundaries" (1977, 84). Such information, then, would feed strategy EX:SUPRASEG.

Other aspects of intonation may also provide useful clues for children. In particular, the characteristic intonation of question tags may also support this heuristic. Adam Clark at 2;3.8 extracted *isn't it*:

Adult: That's an elephant, isn't it?
         What is it?
         (A minute or two later)
Adam: Intit.   (Adam continued to call elephants "intits" for several weeks and was impervious to
                         corrections.)   [Clark 1977, 349]

MacWhinney labels this early strategy "intonational packaging" and suggests that "early forms will take the shape of units which are separately packaged by intonation in the input" (1978, 10). In languages other than English, stress and tone patterns that recur predictably may serve as segmentation clues. Thus, in Hungarian, where main stress always occurs on the first syllable of a word, children rarely make word-segmentation errors (MacWhinney 1976, 389).
Another salient phonological property is the intonation contour itself:

EX:TUNE. An utterance is a speech tune or melody.
Note that this heuristic results in the extraction of the intonation contour itself as an utterance, rather than in the extraction of a segmental stream delimited by that contour, as with the previous heuristic. EX:TUNE is the same as the strategy I have elsewhere called "learning the tune before the words" (Peters 1974), where the "tune" (i.e., the intonational contour) of a particular phrase seems to have been more salient, and hence more memorable, than its segmental phonemes. English phrases that can have such characteristic tunes include *uh-oh! look at that! what's that?* In order to produce carriers for these tunes Minh often mumbled, using nonsense syllables or poor approximations of the adult sequence of segmental phonemes (Peters 1977).

The final phonological property that seems relevant here is rhythmic pattern as manifested in stress, vowel length, and/or number of syllables:

EX:RHYTHM. An utterance is a rhythmic pattern of speech.
There is ample evidence from early language production for the importance of such a rhythm-based heuristic. It is manifested in the appearance of "filler syllables," which seem to be incompletely analyzed syllables that are nevertheless reproduced in a "fuzzy" form that maintains the meter of a particular phrase. Such syllables have been reported in the early productions of children learning English (Bloom 1970; Braine 1976; Peters 1977), Spanish and Cakchiquel (Tolbert 1978), Turkish (Ekmeci 1979; Slobin & Aksu 1980), and Siswati (Kunene 1979).

It should be noted that in any of these examples more than one heuristic could have had a role, although the heuristic it is placed with is probably the most relevant. It is well to remember that these saliency-based heuristics interact with and mutually reinforce each other. Thus an utterance that is at once frequent, often bounded by silence, and possessed of characteristic rhythm and intonation contours is likely to be much more memorable (and thus earlier learned) than an utterance distinguished by only one of these characteristics .

2.2. Factors Affecting The Size Of Extracted Units
In reviewing the literature on early language acquisition one finds that some but not all of children's early language productions (their first "words") are targeted on single adult words. Moreover, this variation can be found both within and across children. A number of terms have been coined to describe the opposing extremes of such variation, extremes that are usually described in terms of constellations of typical observed characteristics. Nelson (1973) has used the terms "Referential" and "Expressive," Peters (1977) has used "Analytic" and "Gestalt," and Horgan (1980) has used "Noun-lovers" and "Noun-leavers." These extremes are generally considered to define the end points of continua along which most children are ranged, very few being clearly at one pole or the other.

What factors in the child's language-learning situation may be responsible for the observed variations in the size of children's first productions - and by inference in the size

of the units they first extract? At least four such factors can be cited: (I) the child's communicative needs, that is, the functions language must serve for her or him, (2) the type of speech the child is exposed to, (3) the type of speech that is expected from the child by her or his caretakers, and (4) the child's individual neurological endowment, which may also underlie any idiosyncrasies of personality.

As we look at each of these factors in turn we must keep in mind that, although each can be described independently, they all operate simultaneously and interact in complex ways. Not only may the picture be somewhat different between any two children, but even within a particular child these factors may combine in different ways in different situations to produce observable variations. Furthermore, this very complexity makes it difficult to disentangle the factors for separate inspection: There is no clear priority and hence no easy place to begin. Bearing this in mind, let us plunge in.

2.2.1. The functions of language for the child
Although the field of linguistics proper focuses on the forms of speech, socio- and ethnolinguistics have been concerned with its uses as well, recognizing that conveying information is only one of many potential functions. Halliday (1973), in particular, has identified a whole range of language functions, including Interpersonal, Regulatory, and Interactional, all of which involve various kinds of social relations. In fact, almost any utterance is directed at achieving more than one communicative goal: "I'm cold" may convey the information that I am indeed cold, while simultaneously requesting that something be done about it (e.g., shutting the window) and indicating that for some social reason I am unwilling to ask my hearer directly to do it.

In spite of such inherent complexity for adults, when children start to learn how the language system works, they may begin by focusing more on one function than on another. Nelson (1973) has proposed that some children decide that their primary need for language is for talking about things (the Referential function of language), whereas other children decide that language is primarily useful for social interaction (the Expressive function). Since Referential children tend to learn names for things and people, their early vocabularies contain many nouns, whereas Expressive children are likelier to learn the verbs, adjectives, and especially the social phrases (such as *hi, thank you, oh boy, all gone*) useful for interacting with others. Note that these are tendencies, not all-or-none phenomena. Thus Nelson found that before the eighteen children she studied reached the fifty-word vocabulary level, "the number of phrases produced by the R[eferential] group during this period ranged from 0 to 5 (mean 2.4) while those of the E[xpressive] group ranged from 6 to 18 (12.6)" (1973, 22).

Halliday, too. has proposed a functional split in early language use between Mathetic and Pragmatic (1975, 17-28). Mathetic language contributes to a child's learning about the environment, whereas Pragmatic language includes the instrumental, regulatory, and expressive-conative uses that contribute to interpersonal relations. This dichotomy seems to correspond rather well to Nelson's Referential versus Expressive.

Furthermore, two strategies of early language production that I have labeled Analytic and Gestalt (Peters 1977) seem to be fostered differentially by the need to use language mathetically (referentially) and the need to use it pragmatically (expressively). An Analytic utterance is "the nice neat one-word utterance" that over time "slowly [increases] . . . in closeness to the adult target" (563); in Gestalt utterances the segmental fidelity is poor, although "the combination of number of syllables, stress, intonation, and such segments as [can] be distinguished [combine] to give a very good impression of sentencehood" (564). My observation was that for Minh, the child in that study, Analytic speech was generally used in referential contexts, such as naming pictures in a book, whereas Gestalt speech tended to be used in more sociable contexts, such as playing with his brother, or in commenting about objects rather than naming them (566). Note that Mathetic/Referential seems to have to do with cognitive development and learning about the structure of the world outside the self. Older children, however, such as those in Wong Fillmore's study of 5- to 7-year-olds acquiring English as their second language (1976), are well beyond this stage of cognitive

development: They do not need to learn about the world, but they do need social interaction; hence they tend first to acquire Interpersonal/ Expressive language.

As Nelson's data suggest, the purpose for which children use their language can affect the kind and size of initial units that they tend to extract from the speech stream. The cognitive (Mathetic) need to label aspects of their daily world (a need that may be fostered by the kind of input speech and expectations about language use that they encounter; see 2.2.2) can motivate them to try to extract such labels from the speech stream. In fact, children soon learn some means of eliciting labels from adults; it may be a point and a grunt, or it may be a formula such as Minh's [as] or [s] ('what's that?'). Children are also often heard to label objects aloud to themselves with no attempt to communicate to others, as if the association of the sound with the object were part of a cognitive process of recognition. Since most labels tend to be single words or tightly tied phrases (e.g., _choochoo train, rocking horse, kittycat_), the result is that Referential or Mathetic-oriented children seem to produce (and presumably extract) more short, Analytic utterances.

On the other hand, the children's need to interact with their environment, both to get others to do things they are unable to do alone (regulatory) and to acknowledge the presence of an important caretaker (interpersonal), will drive them to extract from the speech stream the necessary language for conducting such pragmatic interactions. Since this type of interaction tends to be conducted, at least in English, through multiword formulas or sentences (e.g., _give me that, you do it, nice to see you, I want, that's mine_), the result is that Expressive or Pragmatic-oriented children seem to produce (and presumably extract) proportionally more long, Gestalt utterances.

Although it may be possible to classify most children at the early stages of language learning into dichotomous categories such as those just discussed, such classifications will necessarily be rough and based on some sort of predominant strategy the child seems to be using. But Nelson reminds us that since ' most children learning language learn to use it in a variety of contexts for a variety of purposes, most children will exhibit aspects of both formulaic and analytic approaches in their early language . . . It appears that children can master languages in either way and probably in both at once" (1981, 182-3).

2.2.2. Input speech

Since the speech heard by children constitutes the primary data they have available from which to discover the workings of the language to which they are exposed, it is reasonable to ask whether variations in input speech might be reflected in observable variations in the type of speech first produced by different children, and by inference in the size of units they first extract. Although numerous studies on language input have appeared in the last ten years, the kind of evidence that is relevant to variation is barely beginning to appear. There are two types of reasons for this: theoretical and methodological .

The theoretical reasons have to do with the focus of most input studies, which have generally been concerned either with interaction between mother and child (e.g., how babies younger than I year learn to take conversational turns; see Sachs 1977a; Snow 1977) or with linguistic support for emerging syntax in children 18 months or older (Cross 1977; Garnica 1977; Shatz & Gelman 1977), rather than with extraction of early units (but see Nelson 1973). Even those studies which include children in the crucial (for extraction) 10- to 20-month-old range have been more interested in the effects of input on syntactic measures such as number of noun or verb phrases per utterance than on size and type of units extracted (e.g., Newport, Gleitman, & Gleitman 1977).

One likely cause for the lack of interest in extraction variables is that little variation was noticed, probably owing to methodological factors. Thus the observed mother-child dyads have been predominantly from the middle or upper-middle class in English-speaking communities (United States: Garnica 1977; Newport et al. 1977; Phillips 1973; England: Snow 1977; Australia: Cross 1977); the children have tended to be first-born (Snow; Phillips); and the recording situation has often been either a soundproof laboratory playroom (Garnica; Newport et al.; Phillips) or the child's playroom at home (Cross). There is growing evidence, however, that this homogeneity of culture, social class, birth order, and recording situation

may have conspired to produce homogeneity in the kinds of language used by both mothers and children - a homogeneity that is not generalizable to caretaker-child interactions in other cultures, social classes, or interactive settings.

Nelson was the first researcher to notice these methodological limitations. The group of children she studied consisted of eighteen middle-class American children of whom eleven were first-born and seven were second-born (1973, 60). The parents' educational levels ranged from high school to professional degrees. The children were between 10 and 15 months old at the beginning of her project, and almost all data collection was done in the children's homes. One of her interests was early vocabulary development, and much of the data were collected in diary form by the children's mothers during their normal daily activities. Nelson noted a tendency for first-borns to be Referential (seven to four) and for second-borns to be Expressive (four to three), although this was not statistically significant (60). But when she also took the parents' educational level into account she found that "*all of the firstborn children of the most highly educated families (those with college educations and better) were found in the R[eferential] group*" (61, emphasis hers). She goes on to comment:

As this is the pool from which the vast majority of previous samples employed in the study of early language development has derived, it is an important cautionary finding. It indicates that characteristics thought to be general - even universal - may be confined to a group that differs importantly and systematically from other groups within the general population. [61]

Nelson, then, gave us the first hints that there might be both cultural and familial factors that affect input speech, which in turn might affect a child's route to language.

More recently, three ethnographically based language-acquisition studies, carried out in very different cultural settings, have shown that input speech can indeed be very different from that observed among middle-class English-speaking mothers. These studies are Schieffelin's (1979) study done among the Kaluli of New Guinea, Ochs's (1980) study done in American Samoa, and Heath's (1982, 1983) study of a black community in the Piedmont Carolinas.

Schieffelin found that the Kaluli are very conscious of the language development of their children and deliberately try to teach them to speak by direct instruction, that is, by telling them appropriate things to say in specific circumstances. Such speech, however. is normal adult speech; in fact, the Kaluli find the idea of "baby talk" dismaying. Even young babies are seen as "individuals with intentions, ideas, and identities" (1979, 108), and it is felt to be important for mothers to help these identities emerge, at first by speaking for their infants, then by instructing their toddlers in how to interact in appropriate social and linguistic ways. Although some Kaluli instruction takes place in dyadic situations (when only mother and child are interacting), a much larger portion takes place in triadic situations, in which the mother coaches the child on what to say to a third party. What are the effects of such language on the size of early language units'? Since Schieffelin deliberately chose children who were beginning to put words together, they were all beyond the initial extraction stage. Thus she gives no data on their successes or failures in segmenting adult-sized words out of such speech. One can speculate, however, that this unsimplified type of input might foster extraction strategies different from those encouraged by mainstream American input.

One thing that struck Ochs about caretaker speech in American Samoa was the absence of the expansions of children's utterances so typical of American mothers' speech. After comparing Samoan and American middle-class beliefs about status, rights to speak, and capability to act intentionally, she concludes that Samoan caretakers do not expand their charges' utterances because young children are not considered really capable of communicating intentionally. It is also not appropriate for a (higher-status) caregiver to try to accommodate a (lower-status) child's perspective by offering confirmation checks (in the form of expansions) about what the child might have intended to say. On the contrary, much of caretaker speech consists of providing social instruction for the young child, including elicitation of imitations of people's names and socially appropriate greetings, questions, and so on (1980, 32). In this way Samoan caretaker speech seems similar to that

among the Kaluli, although the Kaluli have a strong language-instruction component in their caretaker talk that seems to be absent among the Samoans (37). Samoan input is also similar to Kaluli input in that it is not marked by "baby talk lexicon, special morphological modifications (diminutives, etc.), simpler syntactic constructions or constructions of reduced length" (36). On the other hand, since much of the caretaking in Samoa is done not by the children's mother but by the older girls of the family, there may be other, as yet unexplored variables in the type of speech heard by Samoan children. These features might make their extraction problem somewhat different from that faced by the American child.

Heath's study (1982, 1983) includes observation of language learning in a small all-black residential community in the Piedmont Carolinas. In this community, which she calls Trackton, adults do not think of children as conversational partners until they are skilled enough to be "seen as realistic sources of information and competent partners in talk" (1982, 114). Since Trackton mothers, like the Kaluli and Samoan mothers, usually have other, more skilled, conversational partners available, they are not limited to talking with toddlers. But neither do they exclude young children from adult talk. Direct communication with infants is primarily nonverbal, since they are carried about all day. Linguistically the children are exposed to streams of talk that continually flow past them. And their first attempts at speaking (between 12 and 24 months) tend to consist of imitations of the ends of overheard phrases or sentences. Heath's examples of such imitations show children picking up in this way chunks that are several morphemes long. It seems, not surprisingly, that given such input they extract, not single words, but longer chunks that are probably intonationally salient as well as intonationally delimited.

In addition to the overall cultural patterns, factors within the family may result in differences in the type of speech heard by the child, leading to variation in initial extraction behavior. For instance, the influence of sibling speech may be very important for second and later children, especially in cultures in which the nuclear family is the rule. This is an area that has hardly been touched; the only study I have seen is Montgomery (1977), which, since it is the study of only one sibling pair, raises more questions than it answers. The questions, however, are intriguing, and more systematic investigation of this area is needed. Montgomery finds that

parental estimations of the advanced level of acquisition of the second child [relative to their estimation of that of their first child at the same age] are not verified in the data. In the analysis of imitation data it is found that the child who imitates . . . the sibling is indeed making a great number of significant and situationally appropriate remarks, thereby giving the illusory impression of advanced competence. [4]

The presence of an older child in the family provides the younger learner with a model of how to interact with others, a model that may well be absent for firstborns. Thus, by 14 months, Minh had learned the "summons-response routine," which was frequently modeled for him by his older brother: He frequently tried this out by calling out "Mommy!" and his mother dutifully and predictably responded with either "What?" or "Huh?" Minh's linguistic development was still so rudimentary that he could not make effective use of this routine to introduce a topic of talk, but he clearly knew that its function was to get his mother's attention.

Older siblings may also provide models for pronoun usage, modeling phrases with incorporated pronouns that younger learners first extract as chunks (e.g., <u>*That's mine*</u>*!,* <u>*I dunno*</u>*,* <u>*I don't wanna*</u>). Clark notes: "The area of pronouns would seem to be one in which first children are at a particular disadvantage. Since they do not have other children playing the same role as themselves with respect to their parents, they do not have the same opportunity for modeling utterances" (1977, 352). And Nelson observes:

The mother who has a 3- or 4-year-old to cope with, as well as a 1- or 2-year-old, will use characteristically different language in interaction with both children than will the mother who has only one child of I or 2 years. A larger percentage of the function of language that the younger sibling hears is likely to be directive and centered around the child's own activities - to be, in effect. pragmatic and expressive. Thus the child is likely to conclude that language is a pragmatic medium that is useful for social control and social exchange, and this conclusion is likely to be shored up by exchanges with siblings. [1981, 181]

Since English does not inflect either nouns or verbs for gender, we have not noticed how the sex of a child can affect the (adult) composition of early extracted units. Hebrew, however, does inflect verbs for gender. Berman (1978) found that her daughter Shelli's first verbs were extracted from the feminine forms that she heard addressed to herself, for example, *kli* (< *tistakli*) 'look! (fem.)', or *shvi* 'sit down (fem.)'. Shelli used these forms when addressing or referring to members of either sex.

Another source of variation in initial extraction behavior has so far escaped notice because the range of interactions during which data have been collected has been so limited. As already noted, the recordings on which many of the input studies were based were made in playroom settings, often in laboratory playrooms on university campuses. During these interactions the mothers were generally encouraged to focus on the speech and/or behavior of their children; often they were not told until after the recordings had been made that the purpose of the study was to look at their own speech (e.g., Garnica 1977; Newport et al. 1977). Such recording conditions conspire to orient middle-class American mothers to try to elicit speech and certain kinds of direction-following behavior from their children. In these dyadic interactions there is much labeling of objects and attributes, along with repetitions and rephrasings of directions. This kind of Mathetic/Referential speech on the parts of the mothers is, on the whole, optimally structured to foster extractions of small units, usually one word long, that are Referential in nature.

But is this a representative sample of the kind of linguistic input that these middle-class American children receive? In the course of their daily lives do they not also participate in other, more social, types of interactions, especially at mealtimes, bedtime, diaper changing, or the arrival and departure of visitors'? Although triadic interactions such as those found among the Kaluli may occur much more rarely in the mainstream English-speaking population, partly because of the relatively small family sizes, there is some evidence that some of these do occur, at least in the context of teaching social skills (Gleason & Weintraub 1978). Systematic efforts to collect such speech have not yet been carried out, though studies of specific routines, such as the trick-or-treat routine at Halloween (Gleason & Weintraub 1976), have been done.

This omission, though partly owing to lack of awareness regarding the potential importance of such routines to the language-acquisition process, is also due to difficulties in collecting these kinds of data. It is much easier and less invasive of the subjects' privacy to bring a mother and child into an acoustically treated laboratory playroom for an hour of recorded interaction than it is to try to record the language used in the regular daily life of the home. Collection of the latter kinds of data requires, at a minimum, that the researcher come to the home and record for periods of several hours in a single day (see, e.g., Bloom 1970; Miller 1979; Ochs 1980; Schieffelin 1979) in order to capture a variety of interactions between caretaker and child, including the comings and goings of any visitors. Even better would be well-designed diary studies supplemented by selective audio recording in which the parent acted as principal data collector.[1] Although it may be more difficult to enforce objectivity in data collection by parents, these may be the only kinds of data that will give a true picture of the range of input that children really receive. And if social routines, which involve Pragmatic/ Expressive speech, foster the extraction of longer units, we are missing half of the picture until we carry out such studies.

2.2.3. Expectations regarding children's speech participation and productions

The child's linguistic environment consists not only of the input speech, but also of a culturally determined set of expectations regarding appropriate language use and acceptable language style. These expectations have not been much discussed until very recently because investigators have tended to conduct language-acquisition research

---

[1] Braunwald & Brislin (1979) give excellent guidelines for designing such studies, and Ferrier (1978) is a nice example of a diary-based study that is able to link child and adult utterances with social and physical contexts.

within their own cultures (for obvious reasons having to do with comprehensibility and access), and hence have not been aware of the ways in which differences in such expectations could affect the process of language development. In particular, these expectations form the unconscious basis of feedback from caretakers to children concerning whether their early vocalizations "count" as language. Such feedback may influence not only the size and types of units extracted but also children's perceptions of appropriate occasions of use. With the rise in interest in ethnographies of speaking, linguists have become more aware of the cultural expectations surrounding adult language use (see, e.g., the studies in Bauman & Sherzer 1974 and in Gumperz & Hymes 1972), but it was not until recently that such ethnographic concerns were specifically taken up by those interested in children's early language (Blount 1972; Heath 1983; Miller 1979; Ochs 1980; Schieffelin 1979). Although there is not yet a great deal of information on how such expectations differ from culture to culture (but see Schieffelin & Eisenberg 1981), it is worth looking in some detail at what currently exists, because of the potential power of such expectations to influence the size of early extractions. We will begin closest to home.

There is growing evidence that there is a single widespread set of expectations regarding early language use that is characteristic of, but not limited to, middle-class American mothers. These expectations, largely summarized from Heath's description of middle-class parents in the Piedmont Carolinas (1983, chap. 6) include the following:

1. That babies are potential conversational partners as soon as they are born (and perhaps even in the womb), although their conversational turns may have to be supplied for them;
2. That systematic modifications of the language addressed to babies, modifications producing what is known as "baby talk" (see Ferguson 1977 for details), facilitate communication with them;
3. That babies' early vocalizations are interpretable as intentional conversational turns with a meaning deducible from the context;
4. That such vocalizations warrant repetition and expansion by the caretaker, who is usually the mother;
5. That it is important, through modeling and contextualization, to teach babies the proper labels for the people and things in their environment;
6. That the production of such labels can and should be encouraged through contextualized interactive question-and-answer routines, such as "What's that? That's an X. That's right. X" (see Ninio & Bruner 1978 for details);
7. That babies' first words are likely to be names, either of people (especially Daddy) or of objects.

To me it seems clear that this set of unconscious assumptions is optimally designed to foster the extraction of units that are exactly one word in length and that have a Mathetic/Referential function. This fostering is primarily effected through the dual mechanisms of provision of labels (6) and expectation of the production of labels (7). Once we have laid these assumptions out for inspection we can see that they underlie much (although not all) of the caretaker talk in middle-class English-speaking families. Berman has observed similar labeling routines in Israeli middle-class families.[2]

Of course, even within a particular culture there are individual variations among mothers in what they expect about their children's early language. Nelson observed that the mothers she studied could be categorized as Accepting or Rejecting on the basis of whether they accepted or rejected the linguistic productions of their children as legitimate speech (1973, 103-4). Some mothers tended to accept as language practically everything their children produced, whereas others tended not to hear their children's vocalizations as real speech. These are the extreme cases; other mothers presumably fell in between, although Nelson felt that she could classify all eighteen of the mothers she studied as Accepting or Rejecting.

Lieven, too, found that two middle-class English mothers whom she observed differed quite noticeably in their manners of interacting linguistically with their toddlers.

---

2 Personal communication.

Kate, a firstborn child, was 18 months old at the start of the study, and Beth, a secondborn, was 20 months old. According to Lieven, "The speech of both mothers to their children showed many of the features which have by now become familiar" (1978, 178). And yet Kate's mother was observed to make more efforts to communicate with her daughter than Beth's mother: She responded to Kate's comprehensible utterances 81 percent of the time (as opposed to 46 percent for Beth's mother); she was more likely to interrupt an ongoing conversation with an adult in order to make such responses; and her responses tended to be geared to extending the communicative interaction rather than simply acknowledging her daughter's utterance. The relevant underlying assumptions here involve the child's role as a legitimate conversational partner and the acceptability of the child's utterances as real speech. It is noteworthy that the children's language was observed to differ markedly: Kate was much like Nelson's Referential children, whereas Beth was more Expressive. And they "appeared to be using language for different ends. Kate talked slowly and coherently about things happening around her and objects in her environment while Beth devoted more time using her speech to try and engage her mother's interest" (178). Thus it seems likely that these mothers' differing expectations about their daughters' language influenced the children's strategies for breaking into the language system.

If we look at studies of language learning in working-class American families we find that, at least in some groups, the underlying assumptions are roughly the same as those just outlined. This is the case for the community of white mill workers in South Carolina described by Heath (1983, chap. 4). Although Miller's (1979) study of three white working-class families in South Baltimore did not start when the children were young enough to be producing their first words, one can find evidence there for assumptions 4, 5, and 6.

On the other hand, Heath's ethnographic study of black mill workers in South Carolina (1983, chap. 3) reveals a radically different set of assumptions, which can be summarized as follows:

1. That children are not potential conversational partners until they have achieved enough skill to get the floor and introduce an intelligible topic into a conversation;
2. That language rarely needs to be addressed *to* babies, although there may be much talk *about* babies in their presence;
3. That babies' early vocalizations are interpretable as noise rather than as language: Smiles and coos are to be rewarded with nuzzles and physical play;
4. Vocalizations that middle-class mothers would interpret as "words" are not worthy of a response designed to further communication because young children have no need to try to communicate their wishes or needs to their knowing caretakers;
5. That there are no unique labels for either people or objects: People may have multiple names or nicknames, each negotiated with particular (sets of) individuals, and similarly, what one calls an object may depend on the context;
6. That children cannot be taught to talk by adults, but rather must "come up" on their own: When they have something to say they will say it;
7. That children, especially boys, will first start to interact linguistically in socially defined situations, often involving challenges to their status as individuals or possessors of goods.

These assumptions are clearly very different from the middle-class set. And if we look at yet other cultures we find an even greater range of expectations. In our discussion of variations in the kinds of input speech that children hear we have already looked briefly at beliefs in two cultures, Kaluli and Samoan, which differ markedly from our own. If we try to extract a set of expectations regarding children's rights to participate in conversations and the nature of their early speech productions from the reports on these cultures (Ochs 1980; Schieffelin 1979), we find numerous differences from either of the sets just presented. Let us look briefly at the range of variation that is found when we broaden our scope in this way.[3]

---

[3] Schieffelin & Eisenberg (1981) review the data on these cultures as well as on several others, for which the data are more meager. I have decided to include here only the Kaluli and Samoan data, and in a few instances inferences from Blount's study of the Luo in Africa (1972), since I have the most details on these cultures and

1. Rights of the child to participate in conversations:
   a. Children are born with such rights: middle-class English-speaking
   b. Children must achieve such rights through the acquisition of sociolinguistic competence: Trackton
   c. Children must achieve such rights through the acquisition of age and status (i.e., children should be seen and not heard): Samoa, Luo
2. Modification of talk to babies:
   a. Talk to babies must be modified in order to promote comprehension: middle-class English-speaking
   b. Talk to babies should not be modified; correct speech should be modeled: Kaluli
   c. Caretakers don't talk to babies: Trackton
3. Meaning of babies' early vocalizations:
   a. They are interpretable: middle-class English-speaking
   b. They are of no consequence: Trackton, Samoa
4. Response to babies' vocalizations:
   a. They should be repeated and expanded when possible: middle-class English-speaking
   b. They should be accepted but not modified: Kaluli?
   c. They should be ignored: Trackton, Samoa?
5. Language skills that need to be explicitly taught:
   a. Labels for people and things: middle-class English-speaking, Luo, Kaluli (to a lesser extent)
   b. Appropriate ways to interact socially: Kaluli, middle-class English-speaking (a limited set of routines only), Samoa (calling-out routines)
   c. People's names: Samoa
   d. None: Trackton
6. Language-teaching mechanisms:
   a. Question-and-answer routines: middle-class English-speaking
   b. Direct instruction to "say X": Kaluli, middle-class English-speaking (to a lesser extent: social routines and names), Samoa (calling-out routines)
   c. Children will talk when they have something to say: Trackton
7. Babies' first "real" language:
   a. Names of people or things: middle-class English-speaking
   b. Specific words: Kaluli.[4]

What is the effect of these differences in expectations on children's early extraction strategies? At the moment it is still hard to tell, since the relevant data are lacking. We have one hint, however, in Heath's observation that Trackton children's early productions are imitations of chunks lifted from the ends of sentences overheard in surrounding adult conversations. It would seem that a Trackton type of environment might foster an Expressive or Gestalt type of early strategy. The ways that adult expectations about children's early language behavior can influence the type and size of the earliest extractions clearly need investigation.

2.2.4. Individual personality and neurology

The fourth and least tangible factor that may affect early extraction strategies pertains to individual personality and neurological differences. Such differences may explain, for instance, why some of Nelson's subjects turned out to be predominantly Referential in spite of having mothers who both produced and expected Expressive

---

since these additions broaden the picture sufficiently without losing us in a morass of detail. The interested reader is urged to consult Schieffelin & Eisenberg and the references therein. It should further be borne in mind that since both the Samoan and Kaluli studies focused on the emergence of syntax, neither has much information on the extraction of early units or on the cultural beliefs surrounding these early extractions.

[4] Kaluli children are not considered to have begun talking until they are heard to use two specific words: *no* 'mother' and *bo* 'breast', regardless of whether they can already use other words (Schieffelin Eisenberg 1981, 14).

speech, and vice versa. And how else can we explain why the firstborn daughter of an academic American couple clearly preferred an Expressive/Gestalt/Noun-leaver strategy (which persisted at least until age 5), whereas their second born son was a Noun-lover (Horgan 1980)? There also seem to be otherwise unexplainable individual differences among young children in their propensity to imitate speech heard around them. For some children imitations occur in as few as 2 percent of the utterances they produce, whereas for others imitations occur in up to 40 percent of their utterances (Snow 1981b, 10-11, tab. 3). Such an individual motivation to imitate could have a noticeable effect on the sizes and types of early extractions.

Individual differences may also be relevant in accounting for some of the individual variation in the learning styles of adult second-language learners. Thus Krashen observes that "some people seem not to utilize a conscious grammar at all, and are apparently immune to the effects of correction." Such a person "relies mostly on his 'feel' for correctness." Other people, however, "monitor" their language use "all the time, and as a result exhibit little fluency" (1978, 10).

Children with learning disabilities might be seen as pronounced examples of individual differences. Normal children are able to use more than one strategy as the situation warrants: Thus Minh seemed to use a Gestalt strategy of picking up long chunks for Expressive speech situations and an Analytic strategy of picking up single words for Referential ones (Peters 1977). Children with learning disabilities differ from "normal" children, not in having access to totally different strategies, but rather in the relative degree to which they can use the same strategies. In particular, certain strategies may be all but denied to certain children. Weeks (1974) describes in detail a child who seemed to have limited access to certain strategies for language learning but who compensated by making fuller use of other strategies.

Of course, not all children with language-related learning disabilities are deficient in the same processing strategies. One suggestive study of dyslexic children found that they could be classed in several distinct subgroups, two of which were labeled "dysphonetic" and "dyseidetic" (Fried et al. 1981). The dysphonetic children "have great difficulty in reading and spelling words phonetically, though they may have a limited sight vocabulary of whole words which they can recognize and read fluently on flash presentation" (15). Dyseidetic children "read laboriously through a process of phonetic analysis and synthesis" (i.e., sounding out words letter by letter), and "they have a poor memory for visual whole-word and letter Gestalts," which leads to "difficulty recognizing words or . . . learning what letters look like" (15). Electroencephalographic measurements revealed neurological processing differences between the two groups. It would seem therefore that for neurological reasons dysphonetic children would have limited or no access to analytic strategies for processing print, whereas dyseidetic children would be unable to make much use of Gestalt strategies. Although these data concern reading, a parallel set of differences may exist for extractive processing of spoken speech.

The existence of large individual neurological differences among adults is consistent with the anatomical evidence summarized in Whitaker and Selnes (1976). For instance, a *fourfold* difference in the size of the area of the brain that processes visual stimuli has been observed among patients for whom this has been mapped (847), and large individual differences in the architecture of the receptive language areas have also been observed (848). Although some work has been done on interhemispheric asymmetries within infants (e.g., Chi, Dooling, & Gilles 1977; Wada, Clark, & Hamm 1975; Witelson & Pallie 1973), little is known about individual anatomical differences across infants. The potential importance of such differences has been acknowledged by Geschwind, who writes, "I have heard it said that the development of language is remarkably uniform across children. I am sure that there are major uniformities, but I suspect that there are many variations on this basic pattern, *determined by individual differences in organization of the brain*" (1980, 306, emphasis mine). If such anatomical differences can indeed be demonstrated among infants, this would constitute further evidence for variable inborn influences on early language acquisition strategies.

These four factors - perception of language function, type of input, type of speech expected, and individual differences - may explain many of the observed strategy differences among children learning their first language. If we examine the diversity in these factors it does not appear remarkable that unit sizes differ and that one child should tend to isolate phrases while another is prone to isolate labels. Our duty as linguists, however, is not to select our data fastidiously from this variation but to develop and use the entire range as evidence for language acquisition processes.