Second Edition

First Language Acquisition

Eve V. Clark

CAMBRIDGE

This page intentionally left blank

First Language Acquisition

Babies are not born talking. They *learn* language, starting right after birth. How does this process take place? When do children master the skills needed to use language successfully? What stages do they go through as they learn to understand others and to talk themselves? This new edition of Eve Clark's best-selling, comprehensive textbook focuses on children's acquisition of a first language, the stages of development they go through, and how they use language as they learn. It follows children from their first sounds and words to the acquisition of adultlike skills in persuading, instructing, and storytelling, whether children are acquiring just one language or two at once. Skilfully integrating extensive data with coverage of current theories and debates, it is an essential guide to studying first language acquisition for courses in linguistics, developmental psychology, and cognitive science.

EVE V. CLARK is the Richard W. Lyman Professor in the Humanities and Professor of Linguistics at Stanford University. Her books include *Psychology and Language* (with H. H. Clark), *The Ontogenesis of Meaning*, *The Acquisition of Romance*, *The Lexicon in Acquisition*, and *Constructions in Acquisition* (with B. F. Kelly). She is an active researcher in the field who works on all aspects of meaning acquisition.

First Language Acquisition

SECOND EDITION

EVE V. CLARK

Stanford University

To Damon Alistair
for all his talk

Contents

Tables, boxes, and figures

Boxes

Figures

Acknowledgements

This book has grown out of my research and teaching on first language acquisition, where, over the years, I have expanded the range of topics to cover something approaching what small children have to tackle as they take on the learning of a first language, together with the many steps they then take on the way to becoming skilled speakers. This book has also evolved over the years in light of repeated discussions with colleagues and students on what we can – and can't – assume about the process of acquisition, and what each assumption buys – or fails to buy us – theoretically.

1   Acquiring language: Issues and questions

Language is quintessentially human. We use spoken language every day, face-to-face, as a means of communication, while written language allows us to record and hold on to our history across generations. Language itself is very complex. It has a sound system that allows us to use numerous distinct words, a vocabulary of some 50,000 to 100,000 terms for many adults, and a series of constructions for relating these words. It allows us to express innumerable ideas, describe events, tell stories, recite poems, buy, sell, or bargain in markets, administer legal systems, make political speeches, and participate in the myriad other activities that make up the societies we live in. Language allows us to coordinate what we do with others, relay information, find out answers, and carry out everyday activities – gossiping, making puns, writing memos, reading newspapers, learning histories, enjoying novels, greeting friends, telling stories, selling cars, reading instructions – the list is unending. Language use calls for an intricate web of skills we usually take for granted. It is an integral part of everyday life that we rely on to convey wants and needs, thoughts, concerns, and plans. Using language seems as natural as breathing or walking.

But babies are not born talking. They *learn* language, starting immediately from birth. What do they learn? They need sounds and words, meanings and construc-tions. They need to know what to use where and when, how to integrate language with other modes of communication, how to make themselves understood and how to understand others. How does this process take place? When do children master the skills needed for using language successfully? What stages do they go through as they learn to understand and talk? Do the languages they learn affect the way they think?

This book focusses on children's acquisition of a first language, the stages they go through, and how they use language as they learn. In this chapter, I take up some of the issues in that process. I outline some of the theoretical approaches in the field and the assumptions they make before turning to the overall plan of the book.

Some issues for acquisition

When children learn a first language, they could build on preexisting notions of what to represent with language as well as prior notions of communica-tion. Or they could start from nothing and discover what is (and isn't) represented

in language. And since languages differ, their acquisition might also be affected by the properties of each language. For example, the type of language could influence the order in which children acquire specific parts of the language and could also make some elements harder or easier to acquire. Their acquisition could also be affected by social interaction and cognitive development. Factors like these could also determine whether language-learners follow the same path, detect and use the same patterns, and make the same inferences about meanings during acquisition.

A tabula rasa?

Do children have to learn everything about language and language use from scratch? Do they start out at birth with John Locke's tabula rasa, or do they come with certain things already pre-wired? Debate over this has led many to draw strict lines between "nature" (any innate capacities and structures children are born with) and "nurture" (what they gain from experience). Biologists would generally argue that this dichotomy is a false one. From conception on, fetal development is shaped by maternal health and nutrition as well as by the fetal cells that are maturing, so to distinguish nature from nurture in development is close to impossible.

Since children are not born speaking, they must learn language. The question then becomes one of what they are born with that is required for this task. Do they come with innate learning mechanisms to get them started? Are such mechanisms general-purpose aids to learning or specific to language alone? What empirical findings could help answer these questions? A related issue is whether children are born with built-in linguistic categories and structures required for learning. Here again, there has been a great deal of debate. Some have proposed that children come with syntactic categories like "noun" or "verb" already wired in, along with certain structural arrays for combining them. The task would then be one of working out what counts as a noun or verb in the speech children hear. Others have argued that children can discover nouns and verbs by looking at all the linguistic contexts each word occurs in. And still others have argued that they can discover nouns and verbs from the kinds of things they designate – nouns are for people, places, and things; verbs for actions. Even if children are born with a learning mechanism dedicated to language, the main proposals have focussed only on syntactic structure. The rest has to be learnt.

In language, children face a particularly intricate task for learning. Compare learning a language to learning how to put on socks and shoes or to brush one's teeth. It is clear that languages demand a lot more. They are highly complex systems whether one considers just the sound system or the vocabulary, or also syntactic constructions and word structure. The structural elements are just half of what has to be learnt; the other half consists of the functions assigned to each element. Learners must master both structure and function to use language.

Languages differ

Languages aren't all cut from the identical pattern, and this makes a difference in acquisition. They differ in the range and combination of sounds they use – for instance, whether they allow only single consonants to begin a syllable (*top*) or also combinations of consonants (*stop*, *trip*); whether they use pure vowels or also diphthongs (combinations of vowels) in syllables (*heat* vs. *height*). They differ in how many word-classes they have. Some have nouns, verbs, adjectives, adverbs, and prepositions (e.g., English and French). Others place "adjectives" in with verbs. Some use prepositions (*in* the boat), some use postpositions (equivalent to *the boat in*), and some add special case endings, usually suffixes, directly onto the locative noun (here, *boat*) to capture the same meaning. Languages also differ in how they indicate who is doing what to whom. Some use case endings on nouns for this (as in German, Finnish, or Latin), and others word order (as in English or Mandarin). A nominative case ending and a first-position noun may do the same job in different languages.

Languages differ in whether word order serves a grammatical purpose (identifying the subject or object, for instance) or a pragmatic one (identifying information as given or as new). They differ in the meanings that are packaged in words, not only in what they have words for (many kinds of camel, in Somali; many kinds of rice, in Thai; many colors, in most Western European languages) but also in just what meaning-combinations are carried by words (whether verbs of motion include information about manner, as in English *walk*, *run*, *stroll*, *trot*, *meander*, or not, as in languages like Spanish or Hebrew that contain fewer such verbs). Languages differ in how they express causation. They may use a lexical verb like *open* to mean 'cause to open' (*he opened the window*), rely on an auxiliary verb combined with a lexical verb, as in French *faire marcher* 'make walk' (*il fait marcher le chien* 'he makes-walk the dog' = 'he walks the dog'), or add an ending to the verb stem itself to make a verb into a causative, as in Turkish or Hindi.

Languages differ in their basic word orders for subject, verb, and object. They may favor SVO or SOV, for example. And they display considerable consistency with the orders of other elements too. In SVO languages, adjectives usually follow their nouns (English is an exception here), and in SOV languages like Japanese they precede them. The same holds for prepositions that precede their nouns in an SVO language like English but follow (and are called postpositions) in an SOV language like Japanese. Relative clauses fill the same positions as adjectives: In SVO languages, they generally follow the nouns they modify, and in SOV languages they precede them. The basic word order in a language is correlated with the order of elements in many other constructions of that language (Greenberg 1963; Hawkins 1988).

When languages combine one clause with another, one clause may be subordinated and introduced by a conjunction indicating whether the relation between the two is temporal (*when*, *before*, *while*), causal (*because*), or conditional (*if*, *unless*). In some, the subordinate clause can follow or precede the main clause, depending on the general flow of information – what's given and what's new. In others, it

may be restricted to a single position relative to the main clause. For example, in Turkish and Japanese, both SOV languages, subordinate clauses must precede the main clause.

Languages are usually consistent both in their basic word order and in the orders favored across a variety of constructions. These statistical universals are important for speaking and listening. The internal consistencies in a language help speakers keep track of what they are listening to and what they are planning to say themselves. They allow predictions about linguistic units and offer predictable frames for the presentation of information. So children need to learn general structural regularities in the language they're acquiring – whether it is an SOV or SVO language, whether relative clauses and adjectives follow or precede the nominals they modify, whether locative phrases are signaled by prepositions or postpositions, and so on. These properties are important because, once speakers have identified them, they can rely on certain assumptions about the kind of information that can come next in an utterance.

Just as languages display consistent structural patterns, they display consistent lexical patterns in the semantic information they bundle together. Some languages combine information about motion and manner of motion, and put information about the path followed elsewhere. The English verb *stroll* conveys 'move in a leisurely manner', while a preposition like *along* marks the path taken in, for example, *stroll along the bank*. Other languages package motion and path together, and put manner elsewhere. The Spanish verb *bajar* conveys 'go/move' plus 'down' and *salir* conveys 'go/move' plus 'out'. To indicate manner of motion, Spanish speakers must add a participle (*corriendo* 'running') or adverb (e.g., *rapidamente* 'quickly') to convey the equivalent of English *run down* (*bajar corriendo* 'go-down running' or *bajar rapidamente* 'go-down fast') (Talmy 1985). Children must learn how their language packages information at word level.

Knowledge of structure and function informs the assumptions speakers make in interpreting what they hear and in choosing how to convey their meaning when they speak. The structures and vocabulary of a language provide choices for speakers. There is no one-to-one mapping of linguistic constructions (and words) to each situation. Instead, speakers must choose how to represent a particular event to someone else. Did Justin chase the dog, or did the dog run away from Justin? Did Sophie come into the house or go into the house? Did Kate teach the children to tie knots, or did the children learn to tie knots from Kate? In each case, the choice of construction and words conveys a particular perspective on the event (Clark 1997). At the same time, the perspectives speakers can take may be limited by what is available in their language.

Complexity for learning

Languages differ in what is easier and what harder to learn. Researchers have distinguished two sources of complexity for learning: *conceptual* and *formal* complexity (e.g., Slobin 1973, 1985b). Conceptual complexity pertains to the

complexity of the ideas being expressed in language. Children probably develop cognitively at about the same rate in similar societies all over the world. This in turn suggests that they should go through stages in cognitive development at the same rate and grasp similar ideas at about the same age. In general, they master simple conceptual distinctions before more complex ones: the notion of more than one (marked by a plural word-ending), say, before notions of truth or beauty, and the notion of an action being finished (marked by a perfective or past tense ending) before the notion of one event being contingent on another (*if X, Y*). In principle, children should master simpler distinctions before more complex ones.

But since languages differ, the same conceptual distinction may be expressed in a variety of forms. One language might opt for a single word-ending for 'more than one' and use this as an invariant form on every noun, much like the *-s* ending for plural in English. Another might make use of ten or more different plural markers depending on the gender of the noun (masculine, feminine, or neuter), the "shape" of the noun (e.g., whether it ends in a consonant or a vowel), its use with a numeral (*five gold rings*) and what numeral (*five*, *ten*, *three hundred*), and so on, much as in Russian or Arabic (see, e.g., Gvozdev 1961; Omar 1973). It should take children longer to learn how to express 'more than one' in these languages than in English. For one thing, there are more forms to learn, and then there are conditions on when to use each one. Differences in formal complexity affect rate of acquisition.

While no one language appears to be easier to learn overall, there are many trade-offs from one language to another in what is easy and what is hard. The plural system for nouns in a language that uses just one ending to mark 'more than one' should be easy. Yet the same language may have an elaborate system of verb tenses and verb forms in each tense, which makes verbs hard to learn. Children may find some aspects of a language easier to master than others, and children exposed to different languages may well learn at different rates on equivalent parts of the system. To find out, we need to establish what's hard and what's easy in acquisition for each language.

Social dimensions

Language acquisition takes place in mid conversation. Adults and children talk to each other; adults expect children to respond to requests and comments, and to indicate to their interlocutors what they are interested in as well as their needs and wants. When adults talk to children, they directly or indirectly offer them extensive information about their language. They set up both tacit and explicit expectations for when children should talk, what they should say, when and how they should respond to adult utterances; what counts as a turn in conversation, when (and when not) to take a turn; and what counts as an appropriate contribution in the ongoing exchange (Berko Gleason 1988). In the course of conversation, adults use the conventional words for objects and actions. This way, they provide words for whole arenas of experience – food, clothing, toys,

pets, vehicles, birds, mammals, plants, gardens, farms, the seaside, mountain slopes, and many more. They also offer information about how words within a domain are related (Clark & Wong 2002).

Conversation demands that its participants attend to each other and to whatever is being talked about. This means keeping track of what others know at each point in the conversation. The participants share common ground and add to it with each utterance. Both joint attention and the updating of common ground play a role in acquisition (Clark 2002b). In learning to participate in conversations, children learn more of their language and more about how to use it (Snow 1978). And in tuning in to a language, they tune in to those distinctions that are obligatory; they come to assume distinctions that are *always* encoded in that language but not necessarily in others. They learn to think – and plan – for speaking in that language (Slobin 1996).

Conversation provides a forum for using language. It displays language embedded in larger systems for communication and so should present children with critical material for making sense of language as they try to understand others and make themselves understood. Conversational exchanges between children and adults should also be a forum for learning to become a member of the society and the culture. From birth on, the exchanges children participate in attune them to the language around them. This holds as much for sound patterns as for words or for constructions used to convey temporal and causal relations among events; as much for intonation contours and tone of voice (with positive or negative affect) as for details of constructing words from roots and affixes.

Understanding in conversation may depend as much on what is not said as on what is said. Knowing some of the elements of a language doesn't necessarily allow one to interpret utterances appropriately. One has to learn the conventions on use. For example, the request in English *Can you open the door?* is both a question about ability (*can*) and a request for someone to perform the action of opening. The context of use then determines how the addressee should construe it. What counts as a request or as an assertion and the range of forms that can be used depend on the conventions of the speech community. (These are not necessarily the same even in communities using the same language.) Construals also depend on the inferences that are licensed in context.

How do children learn linguistic conventions? For instance, the expected response to a question can depend on both the context and speaker. If a speaker repeats with question intonation what a child has just said, this conveys that the adult considers what the child said to be wrong. In everyday conversation, this typically leads the original speaker to offer some alternative. But in the classroom, teachers may question what children say to check on whether they really know, and this calls instead for the child to repeat the original utterance, not change it (Siegal 1997).

Language use is not uniform; it depends on who one speaks to. In most communities, people speak to family members and friends differently from strangers; they distinguish formal from informal speech (e.g., with *vous* vs. *tu*); and they

use a range of polite forms that differ in terms of address (*Ms. Pipon* vs. *Sophie*), word-choices (*that policeman* vs. *the cop*), and syntactic constructions (*Come here* vs. *Could you come here?*), depending on the language and addressee. Learning what the conventions are, the "rules of use" for different occasions, takes time.

Language is not an autonomous system for communication. It is embedded in and supplemented by gesture, gaze, stance, facial expression, and voice quality in the full array of options people can use for communicating. In learning language, children may first rely on nonlinguistic options, both in their initial understanding and in their own early use. They might understand affect first from adult voice quality and gesture, and infer the locus of attention from adult gaze or stance before they understand that words pick out referents. And they might rely on iconic gestures referring to or anticipating reference to things later named with words. Adults may draw children *in* to language by leaning on nonlinguistic means to signal affect or to direct attention. They may even indicate to young children how things work at first through gestures rather than words.

Cognitive dimensions

What do children know by the time they start talking at age one? They have already had about twelve months of perceptual and conceptual development. They are adept at perceiving similarities, identifying objects and actions, recognizing faces, sorting like with like. They can orient objects and know where they are kept and how they are used (spoons, cups, bowls, bottle tops; shoes, socks, mittens; balls, dolls, soft toys, books; blankets, chairs, staircases). They know a good deal about their surroundings, about Euclidean space (up vs. down, back [not visible] vs. front [visible], side to side) and topological space (inside vs. outside, contained, attached, supported). They display memory for objects (persisting in looking for keys that have been covered with a cloth); they use "tools" (enlisting adult aid to get a box open); and they make use of pretense in play (moving a block while making car noises). In summary, they are setting up representations of what they see and know. They make use of these for recognition and recall, summoning them first with gestures and reenactments of events, and later with words (e.g., Piaget 1952; Werner & Kaplan 1963; see also H. Clark 1973).

Do children make use of this perceptual and conceptual knowledge as they acquire language? The answer has to be yes. When they learn to speak, they represent their experiences in words. They also draw on conceptual knowledge and its organization as they work out the meanings of new words and constructions. This is a major source of hypotheses about word meanings. Children use words to pick out categories of objects, whether "dog" or "Dalmatian," "pet" or "pest." These categories may be at different levels (compare "dog" to "Dalmatian" [a kind of dog]), or they can be orthogonal to each other (compare "dog" to "pet" or "guard"). Children can use words with these meanings to pick out the same object from different perspectives. They can use other words to pick out actions, where their choices depend on the number of participants, the effects, the manner of acting, and

the location or direction involved (compare throwing a ball, opening a door, drinking milk, pushing someone on a swing, walking, sitting down, swimming, and riding a bicycle). Children can also assign words to pick out relations in space (compare putting keys in a box, hanging a picture above the head of a bed, climbing down a ladder, sitting beside the fire, crawling across the floor, or looking at a lid on a box, at tiles above the sink, or at a screen in front of the fire). One issue for language acquisition is how children find out which meanings there are words for; another is just how they map each meaning to the right word.

How do children form conceptual categories in the first place? They start out, it seems, with the ability to group things by how similar they are. These early groupings are also influenced by perceptual Gestalts that highlight "figures" against "grounds." Anything that moves stands out against its background and so is the figure. And when objects move, they move as a whole, so whole objects are more salient than any one part. Once children have represented an object-type, they can go on to attend to the actions and relations that link it to other things around it. These kinds of conceptual organization provide a starting point for what might also be represented in language.

Early conceptual organization also offers clues to how children might learn language. They must be able to use prior experience to recognize when objects or events recur. They need to set up representations of what they see, hear, touch, and taste so that they can recognize recurrences. Without such representations in memory, they couldn't categorize or organize experience. To do this, children must be able to detect similarity or degrees of similarity, a capacity that appears fundamental for all learning.

Learners and learning

Learners can be conservative or bold, or somewhere in between. When children learn language, they could go step by step, one form at a time, waiting for evidence from adult speech and rarely going beyond it – *go*, *run*, *fall*, *fell*, *cat*, *cats*, *feet*. They could generalize from a few forms to new instances – from *jump/jumped* to *run/runned*, from *cat/cats* to *man/mans*. They could go item by item then make some limited generalizations, with different children following different paths. Or they could generalize broadly, acting as if all of language is orderly and rule governed (it isn't), and so regularize many irregular forms (e.g., *bringed*, *sitted*, *goed*, *foots*, *sheeps*, *mouses*).

Take the plural *-s* in English. It has three variants depending on the final sound of the stem, as in *cat/cats* [-s], *dog/dogs* [-z], and *horse/horses* [-ɪz]. This is the regular plural form that appears on most nouns in English. It could be learnt by *rote*, with children adding one item at a time as they hear it. Their first version of a word could be singular or plural, depending on what they happen to hear first. So they might learn *cat* and then *cats*; *stairs* then *stair*; *dog* then *dogs*. Rote learning depends on children hearing each form so they gradually fill in the paradigm of singular and plural for each word. Rote learning should preclude errors like *mans*

for the plural of *man* or *teeths* for the plural of *tooth*. It should also preclude children treating words like *house* and *purse* as if they were already plural. Yet children make both types of errors.

Suppose instead that children learn a few forms by rote and use those as models for deciding on the plural forms for new words: Because of *cat–cats*, the plural of *rat* should be *rats*. Here children would be relying on *analogy* (Gentner & Medina 1998), using information about similar words (similar in, say, sound or meaning or both) in deciding what the plural (or singular) should be. Analogy can start from any point, with children choosing a regular or an irregular form. For instance, analogy from *dog–dogs* applied to *cat* and *sheep* yields *cats* and *sheeps*. Analogy from an irregular word (e.g., *foot*, *child*) runs into problems.

Children might instead consider all the forms accumulated so far and abstract a *rule* for the plural (Pinker 1999). This could be stated as "Add *-s* to nouns to form the plural." When the words are regular, children succeed in producing the correct forms; when they aren't, they overregularize. Just as for analogy, rules fail for irregular words. The rule applied to words like *foot*, *child*, or *mouse* does not result in the conventional *feet*, *children*, and *mice*. These irregular words either require additional special rules or rote learning of each adult form.

Both analogy and rule work by adding a word-ending to the existing word. Children start with a source word, add something, and produce a new form. An alternative is to start from the goal – what the plural form should sound like – and adjust the singular word until it fits. Here children could use a *schema* or *template* for the plural (Bybee & Slobin 1982). The schema could be characterized as requiring a form ending in *-s*, roughly, PLURAL = [word + *s*]. If a word fits this schema (it already ends in *-s*), no change is required; if it doesn't, then the word must be adjusted until it does (by adding *-s*). The schema approach accounts for the same regular forms as the analogy and rule approaches do, and it also accounts for why children fail to add a plural ending to nouns like *horse* or *rose*: They end in an *-s* sound and so already fit the schema for plural.

Do children depend on rote, analogy, rule, or schema? Which account best captures what they do with the regularities they detect in language? The answer depends on careful analysis of the forms children produce: what they get right and what they get wrong. One factor is the identification of recurring patterns and their frequency. Children hear instances of some nouns and verbs more frequently than others (*man* occurs many more times than *field*, and *put* more often than *yell*). This is token-frequency. They also hear some types of nouns and verbs more often than others: There are many more regular nouns (e.g., *book/books*, *cat/cats*, *chair/chairs*) than irregular nouns (e.g., *foot/feet*, *man/men*, *mouse/mice*) in English. The same goes for verbs: Regular verbs (e.g., *walk/walked*, *open/opened*, *jump/jumped*) far outnumber irregular ones (e.g., *go/went*, *bring/brought*, *fall/fell*). To what extent does this token- or type-frequency play a role in children's generalizations?

Researchers agree that children must learn both sound systems and vocabulary. (How they learn them is another matter.) Sound systems are specific to each

language, and children must learn the one they are exposed to (Jusczyk 1997; Vihman 1996). And vocabulary presents a formidable challenge. Adults know somewhere between 50,000 and 100,000 distinct words, so the learning required here is extensive (Bloom 2000; Clark 1993). There is much less agreement about the learning of syntactic constructions. Do children rely on innate knowledge for these or do they learn them as they do words? The arguments for innateness have hinged largely on the putative difficulty of learning syntactic constructions from child-directed speech. Researchers have pointed to the ungrammaticality of adult-to-adult speech and also argued that some constructions are either absent or so rare as to make them unlearnable. If children acquire them anyway, they must be relying on some built-in knowledge. Both premises here are in dispute – that child-directed speech is ungrammatical and that certain structures are unavailable in that speech.

   What role do children play in learning? They could be passive recipients of the language directed to them, simply absorbing whatever they hear, or they could play an active role, selecting and generalizing about whatever they have taken in so far. To what extent are children miniature scientists, testing hypotheses and checking up on what they know about particular words or constructions? Do they detect patterns and apply them to new cases? Do they make inferences about possible meanings and make use of them in later word use? Overall, the role that *children* play provides critical information about how (and what) they learn at each stage and about the learning mechanisms they rely on.

Product versus process

      Some approaches to language acquisition focus on the *product* – the end state to be achieved – rather than on the *process*. This distinction tends to capture one difference between linguistic and psycholinguistic approaches to acquisition. Linguists tend to focus on the product, for instance, what a relative clause looks like, laid out on the table for analysis. In contrast, the psycholinguist is more concerned with when the speaker needs a relative clause, how he accesses the pertinent structure, the phrases, words, syllables, and sounds, and then produces the utterance itself piece by piece. This has led to differences in emphasis, with linguistic approaches focussing more on the adultlike nature of children's knowledge while psychological ones have focussed more on the changes that occur during development.

   One linguistic approach known as parameter-setting proposes that children start out with default settings for *parameters* that capture all the dimensions that distinguish among languages. For instance, languages differ on whether they require subjects to be marked by a pronoun where there isn't a noun subject present. (Where they don't, languages typically mark person [e.g., *I*, *you*, *he*] and number [singular or plural] with endings on the verb, as in Italian.) This is called the Pro-drop parameter, and researchers have assumed that the default value is to drop pronoun subjects (much as in Italian or Spanish). Each parameter has

a start-up setting (the default) and children begin there, regardless of the language to be acquired. Then, at a certain point in development, they identify the actual parameter-setting for that language (it is not clear what the critical data are) and from then on make adultlike use of the pertinent forms. What happens before a parameter is set is of scant interest. The main concern is with the parameters themselves, the values for each, and when the correct setting for each is triggered. Setting parameters is regarded by some as something that happens automatically when children reach the right age and stage of development. This leads researchers to ignore everything that happens before a parameter is set (e.g., Borer & Wexler 1987; Radford 1990). Children's errors prior to adultlike use and any continuity in their attempts to convey a particular meaning are simply not relevant.

Other approaches regard *continuity* of expression and *function* as critical clues to tracing the path children follow as they acquire language. This holds for most processing approaches. For example, they may identify a particular conceptual distinction and then trace its expression by children as they learn more about the conventions of a particular language. Take the notion of plurality, more than one. Children acquiring English often start out by using a word like *more* or a numeral like *two* to express this notion, as in *more shoe*, *two cup*. Only after that do they learn to add the plural ending (*shoes*, *cups*). The earlier expressions for plurality show that children have grasped the notion but haven't yet worked out how to express it in English. This comes back to the distinction between conceptual and formal complexity. Children may have acquired the pertinent concept (here, plurality) but not the forms that are conventional for its expression.

Processing approaches have also focussed on what children do at one stage compared to the next. One approach has been to look at where children start, what they attend to first, and what they change in their language as they get older. Their preferences and the changes they make can be captured as processing strategies or operating principles. For example, in producing words, children focus on the core word (the stem) first and on getting the initial sounds right. This strategy can be represented as "Pay attention to the beginnings of words." It helps others recognize the words children are trying to say. Their next move is to start producing word-endings (like the plural, say): "Pay attention to the ends of words." But now they need to attend to the range of meanings conveyed by word-endings, so another strategy might be to look for endings that have a stable, identifiable meaning and to use those whenever needed.

Researchers have looked for consistencies in how children interpret and produce words from the earliest stages on and from those patterns have derived the strategies children seem to apply (e.g., Slobin 1985b). This approach relies on looking at both what children get right and what they get wrong. Sometimes they fail to produce a form altogether (*I throw ball*, without *a* or *the* before *ball*); at other times, they apply a form incorrectly (*bringed*, *foots*). This approach is concerned both with learning and with how changes come about.

Processing approaches take account of the dynamic nature of conversation. Speakers interact with each other. They don't produce isolated sentences that stand

on their own. Once someone has mentioned *Kate*, for example, the next speaker will use *she* (not *Kate*) to refer to her again. Or, once someone has asked Rod whether he wants lasagna, he can answer *Just a little*, or *Yes please*. What these utterances refer to requires that we know that there was a prior offer, *Would you like some lasagna?* Without that, we can't give a full interpretation to *Yes please*. What someone says depends critically on what someone else has just said and often can't be interpreted without a whole sequence of contributions to the conversation. Imagine recording a conversation and then transcribing what only one of the speakers said. It quickly becomes difficult or impossible to interpret what that person means. In fact, utterances depend on both conversational and physical context for interpretation (H. Clark 1996). This should hold even more strongly for young children whose utterances may consist of only one or two words.

The goal of acquisition

The goal is to become a member of a community of speakers. This entails learning all the elements of a language, both structure and usage. Children need to learn the sound system, the *phonology.* This in turn means learning which sounds belong (sound segments like **p**, **b**, **t**, **d**, **s**, **z**, **a**, **i**, **u**, **e**), which sequences of sounds are legal in syllables and words (phonotactic constraints, e.g., **dr**ip but not **dl**ip in English), stress patterns on words (e.g., *electric* vs. *electricity*), tone on words in a language like Mandarin or Hausa, and the intonation contours in sentences that distinguish a question from a statement (e.g., *Alan is coming at six o'clock?* vs. *Alan is coming at six o'clock*).

They need to learn about the structure of words, their *morphology*: whether they are made up of one syllable, two, or many (compare *pop, slipper, alligator*), along with their meanings. Words can be complex and made up of several building blocks, sometimes with suffixes or prefixes added to root forms (e.g., *write/writer*, *saddle/unsaddle*, *push-chair*, *sun-rise*, *house-builder*, *complexify*, *physicist*). These building blocks also allow for the construction of new words to express new meanings, meanings for which there is no existing conventional form. Words may form *paradigms*, groups that display regular alternations to mark particular meanings. In some languages, nouns can be singular or plural, for example (English *cat/cats*, *chair/chairs*, *horse/horses*), but not all of them belong to regular paradigms (English *mouse/mice* or *child/children*). Nouns may also have suffixes that show whether they have the role of subject (e.g., *The man was running*), object (e.g., *The dog chased the man*), indirect object (e.g., *The boy gave the book to the man*), and so on, as in German, Greek, or Finnish. These case endings, like plural endings in English, are generally fairly regular, with the same form used on many different nouns. There may be several plural endings for different sets of nouns (e.g., masculine, feminine, neuter; or common and neuter) and therefore several regular paradigms. Verbs may belong to many paradigms too, each one marking tenses differently, for example. In each instance, noun and verb endings add modifications to the basic meaning of the roots or stems.

Speakers don't use just one word at a time. They combine them, and again the possible sequences of words in a language have to be learnt. This is the *syntax*. Just as with sounds, some sequences are legal, others not. In English, adjectives precede the nouns they modify (e.g., *the green vine*, not **the vine green*), articles like *a* or *the* and demonstratives like *that* also go before their nouns (e.g., *the whistle*, *that rosebush*). Relative clauses follow their nouns (e.g., *The wallaby that was hopping across the path was a female*). Subordinate clauses introduced by conjunctions like *if, because*, or *when* in English can be placed before or after main clauses (e.g., *When the bell rang*, *all the children came inside*, or *All the children came inside when the bell rang*), but in Turkish or Japanese, for instance, such clauses must precede the main clause. Some constructions allow a number of different nouns and verbs to be used in them; others may be very restricted. Just as with sounds and words, children have to learn what the possibilities are.

Language is used to convey *meaning*. Words, suffixes, and prefixes all carry meanings that are *conventional* (Lewis 1969). The speech community relies on all its members agreeing that *ball* means 'ball', *throw* means 'throw', and *sand* means 'sand'. These conventions are what make languages work. Without agreements about meanings, one couldn't rely on the fact that the next time someone uses *sand*, say, people hearing the word will still interpret it in the same way. Conventions are critical in language use. They govern both word meanings and construction meanings. In learning a language, children must learn the conventions for that community.

Languages work in large part because they don't use needless duplication. Each conventional word differs from all its neighbors. Each word reflects a choice made by the speaker to convey one meaning rather than another and so *contrasts* with all the others (Clark 1990). If speakers wish to convey a meaning for which there is no conventional word, they can construct a new one to carry that meaning. This new word then contrasts with any previously established ones. For example, the verb *to skateboard* was introduced along with skateboards themselves to talk about a new method of travel. This verb immediately contrasted with all existing verbs for other means of moving (*to bicycle*, *to sled*, *to ski*, *to roller-skate*, etc.) (see Clark & Clark 1979). Language, and especially its vocabulary (the lexicon), is not static. Speakers coin new words as society changes and adds new inventions and new technologies. But each new word is accepted only if its meaning contrasts with the meanings of existing words. Conventionality and contrast are powerful pragmatic principles governing language use (Clark 1993).

Knowing what the conventions are for the elements of a language and knowing how to use them are two different things. Children must learn how each word and construction can be used to convey their intentions. They learn how to make assertions (*That's a tadpole*), requests (*Can you mend my yoyo?*), and promises (*I'll mow the lawn tomorrow*) (Austin 1962; Levinson 1983). They learn what counts as polite (*Pick up the other one!* vs. *Could you bring in the other box?*), and how polite to be on each occasion. They learn how to give directions and

explanations, and how to tell stories. In summary, they need to learn to use language effectively, whatever the genre, whoever the addressee, and whatever the goal.

Stages in acquisition

Infants don't produce their first words until age one or later, but by three or four, they can talk quite fluently about some topics. This development is one we take as much for granted as the infant's transition from lying supine in the first few months to walking and running around by age one to two. Learning to talk is more complicated than learning to walk. Talking plays a major role in social communication and demands a grasp of all the local conventions of use in each speech community. Language use is an integral part of communication; it goes along with gesture, gaze, and other nonlinguistic means used to convey attitude and affect as well as speaker intentions.

As children learn to talk, they go through a series of stages, beginning with infancy, when they are unable to converse and do not yet understand any language. They go from babbling at seven to ten months old, to producing their first recognizable words six to twelve months later. Then, within a few months, they combine words and gestures, and produce their first word combinations around age two. This is followed by the production of ever more complex, adultlike utterances, as they become active participants in conversation, taking turns and making appropriate contributions. They begin to use language for a larger array of functions – telling stories, explaining how a toy works, persuading a friend to do something, or giving someone directions for how to get somewhere. Between age one and age six, children acquire extensive skills in using language and can sound quite adultlike much of the time. By around age ten to twelve, they have mastered many complex constructions, a good deal more vocabulary, and many uses of language.

Comprehension, throughout this process, tends to be far ahead of production. Children understand many words long before they can produce them, and this asymmetry between comprehension and production is lifelong: Consider the number of dialects adults can understand without being able to produce more than two or three at most. For a second language, consider how much better people are at understanding than at speaking. The same holds true for a first language: Comprehension remains ahead of production, but once production reaches a certain level, speakers tend to no longer notice any mismatch (yet it is still there). At the same time, mismatches play an important role in the process of acquisition: Children's representations for comprehension provide targets for what their own production should sound like.

Is there continuity over stages? Do children try to express similar notions at successive points in development – whether issuing one word at a time, longer word combinations, or adultlike phrases? How much consistency is there in the stages children go through as they learn the same language? How much for

children learning different languages? Do children from different social classes go through the same stages provided they are learning the same language? Are they all exposed to the same amount and same range of child-directed speech?

Why study acquisition?

In the late 19th century, the burgeoning study of child development emphasized language, and many researchers kept extensive diaries of their children's development, including language (e.g., Ament 1899; Baudouin de Courtenay 1974; Compayré 1896; Lindner 1898; Major 1906; Preyer 1882; Ronjat 1913; Stern & Stern 1928; Sully 1896; Taine 1870; see also Campbell 2006). Because researchers lacked tools for preserving their observations, these records vary in quality. There was no audio- or videotape to record what happened and no International Phonetic Alphabet to help note children's exact pronunciations. Some, like Clara Stern and William Stern, who kept a detailed diary, though, raised many issues that are still critical in the twenty-first century. These observational studies were followed by extensive records of children's vocabularies in terms of size and content at different ages. In the 1930s and 1940s in the United States, the emphasis remained on vocabulary size and sentence length, with little analysis of structure and no analysis of conversational skill.

In the 1960s, under Noam Chomsky's influence in linguistic theory, researchers renewed their interest in how children acquired language. Chomsky himself argued that children must rely on certain innate structures and mechanisms, specific to language, because it would be impossible for them to learn from adult speech alone (but see Chapter 2). These claims became embedded in the Chomskyan approach, although few of his students did empirical research on language acquisition in children. Among psychologists who took up the challenge of studying language acquisition directly was Roger W. Brown. He in turn drew many of his students as well as others into the field during the 1960s and 1970s, made major contributions himself, and has had a lasting impact.

Initially, many studies of language acquisition were undertaken to assess the psychological reality of a linguistic proposal or to test the predictions of linguistic theory against acquisition data. And here several problems arose immediately. First, linguistic theory for the most part is a theory about product and not process, so it was unclear what the predictions should be. Even when these appeared fairly clear, there was frequent disagreement on how to interpret findings inconsistent with the current linguistic theory, with linguists commonly dismissing acquisition data as irrelevant and, therefore, as no test for the theory. Second, linguistic theories displaced each other with some rapidity, so theoretical claims became even harder to evaluate. These factors led to some divergence in approach, with much of the research on language acquisition being carried out at some distance from theoretical claims in linguistics. This encouraged the development of other approaches to acquisition and may have led researchers to ask broader questions than they might have done otherwise.

Some of the current issues are still those that dominated debates about language acquisition after the publication of Chomsky's *Aspects of the theory of syntax* in

1965. One of these is whether there is a mechanism for acquisition specialized for language alone, independent of other cognitive skills. This claim has generally been accompanied by the claim that some knowledge about language is also innate, with syntactic categories (word-classes like noun and verb) and basic syntactic structure (subject and predicate, along with other basic grammatical relations, for example) being the prime candidates. This in turn has led to discussion of how much of language is learnable and under what conditions (where the focus has again usually been on syntax alone); whether there is a critical period for language learning, after which humans can no longer learn a language, in much the same way that goslings can no longer imprint on a mother goose or white-crowned sparrows can no longer learn the songs characteristic of their species; and how children learn to correct any errors they make, given the supposed absence of corrective reactions from adults.

The problem with many of these debates lies in the virtual absence of empirical findings and testable hypotheses. The premises have all too often been regarded as facts, and the arguments have raged from there on in. What are needed are testable hypotheses and analyses of pertinent data by the researchers making the claims. Ideally, their questions should yield answers from actual findings on acquisition. These debates, largely carried on in the pages of linguistics books and journals, have ranged over nature versus nurture, innateness (what's innate and "special" about human language) versus learning (what might be learnt, or not, from child-directed speech), and, more recently, the social versus cognitive properties of language as a tool for communication or a system for the representation of knowledge.

My own emphasis is on the social setting of acquisition combined with the cognitive foundations children can build on. So I view both social and cognitive development as critical to acquisition. Since it remains unclear how much of language is innate or whether any specialized learning mechanisms subserve it, my stance on this is a conservative one. I prefer to see how much one can account for on more general grounds first. The emphasis here is therefore on how (and how much) children can learn from adult usage, including specially tailored child-directed speech. I also look at evidence for early generalizations versus initially piecemeal acquisition of constructions with specific verbs and other lexical items. I place considerable emphasis on the developmental processes required in learning a language from the first words on and none on arguing that children know (nearly) everything from the start. As a result, I emphasize continuity in development – continuity in the meanings they express as they move from one word at a time to adultlike utterances for conveying their needs, their interests, their attitudes, and their thoughts.

The plan of this book

Language is social. For language to work, speakers must ensure joint attention with their addressees and then make every effort to achieve and maintain

common ground in each exchange. Its successful use depends on collaboration and cooperation among speakers. In this book, I start from that premise as I follow different themes through the process of acquisition. These themes include the roles of social and cognitive factors in language acquisition; the extent to which children learn different languages differently – how the course they follow is shaped by properties of each language; the increasing complexity of the expressions acquired with age; the stability children display in their order of acquisition for meanings and structures within a language; the role of common ground and the flow of information; the speaker's choice of perspective marked through words and constructions; and the importance of pragmatic factors in the acquisition and use of language, and what might constitute plausible mechanisms for acquisition.

Language is an elaborate resource for communication. It is complemented by various nonlinguistic resources – gesture, gaze, facial expression, bodily stance and orientation – that, together with language, make up the general repertoire people draw on to communicate. Language itself depends on a complex set of conventions on the meanings and uses of words and constructions. Without these conventions, speakers couldn't be sure that words, for instance, had the same meaning from one occasion to the next or from one speaker to the next. So, in learning a language, children need to learn both its conventions and how to apply them. The goal in acquisition is mastery of the language in use around them, so analyses of acquisition must be based on the language children hear. This *use-based approach* to acquisition takes actual usage as the target rather than any idealization of language. The words children hear and the constructions those words appear in are drawn from local patterns of usage in the speech community. The social setting where children are exposed to a first language is critical; this is where they hear their language used. This is the material they must learn to recognize, analyze, understand, and produce themselves.

To study acquisition, then, requires that we look at how children use language, what they have learnt about carrying on a conversation – for instance, taking turns, uttering different speech acts, taking account of what the addressee knows, and connecting new information to what has already been given. This approach encompasses both the acquisition of structure (forms and their meanings) and function (what forms can be used for and how they are deployed for each purpose). The same use-based approach must apply where children acquire more than one dialect or more than one language at a time: learning two (or more) at once, and when to use each, again depends on the usage within the community.

This book is divided into four parts. In the first (Chapters 2–6), I begin by looking at children's conversations with adults and the information adults offer them about language use (Chapter 2). Next I turn to how children analyze the speech stream to recognize words (Chapter 3) and then review the content of children's early words – the kinds of meanings they express – and how they learn to pronounce them (Chapters 4–5). I end with how children map meanings onto words (Chapter 6). The emphasis is on how children get started and their earliest uses of language.

In Part II (Chapters 7–11), I focus on children's acquisition of structure. They learn first to combine two or more words in a single utterance (Chapter 7) and modify each word with appropriate endings (Chapter 8). They add complexity to what they say in two ways: (a) by elaborating the information inside clauses (Chapter 9) and (b) by combining two or more clauses (Chapter 10). In each case, children advance from rudimentary expressions of meanings to more elaborate ones that use conventional adult forms. Lastly, I look at how children coin words when they don't have any ready-made for the meanings they wish to convey (Chapter 11). The emphasis here is on how children acquire the adult forms for their meanings. With both constructions and coinages, they gradually build up more elaborate communicative options.

In Part III, I turn to the social skills children need. They take part in conversations quite early, but learning what to say when is complicated, and getting the timing right for taking turns is also hard (Chapter 12). On top of that, learning how to be polite, to be persuasive, to give instructions, or to tell stories all take added skill (Chapter 13). Finally, children exposed to two languages from the beginning have two systems to learn, and are also continually faced with the decision of how to talk – which language (or which dialect) to use. These choices, just as in the case of one language, depend on the addressee, setting, and topic (Chapter 14). All these social dimensions of language acquisition complement the structural ones. Children have to master both to become identified as speakers from a particular community.

In Part IV, I take up biological specialization for language and where in the brain language is processed (Chapter 15). I then review the kinds of mechanisms needed for the acquisition of a system as complex as language, demanding a wide range of skills for use (Chapter 16).

Throughout, I draw on data from a range of languages to underline both similarities in the analyses children do and differences in how speakers do things from one language to another, and, for both cases, the effects this can have on acquisition. I draw extensively on the diary study I kept of my son from birth to age six to illustrate some facets of language development described here. These observations are supplemented by other longitudinal records and by experimental data on the comprehension and production of specific constructions. I also draw extensively on other published findings and on data from the CHILDES Archive, a collection of transcripts from different researchers (MacWhinney & Snow 1985, 1990). Despite a plethora of studies since the 1960s, there are still many gaps in what we know about acquisition, even for well-studied languages, and there are still too few language-types included among those for which we do have data (Slobin 1985a, 1992, 1997). I hope the present overview will inspire readers to ask further questions, look at as-yet unstudied languages, and take up new questions about the many intriguing puzzles of acquisition.

Getting started

It seems to us that a mother in expanding speech may be teaching more than grammar; she may be teaching something like a world-view.

Roger Brown & Ursula Bellugi 1964

[S]peech skills have a tremendous potential for assisting the formation of non-linguistic categories. The total list of such categories that a child must learn is a cognitive inventory of his culture. Speech, therefore, is the principal instrument of cognitive socialization.

Roger Brown 1958b

The chapters that follow look at the setting in which infants are first exposed to language and in which they take the first steps towards making use of it. This setting is a social one, where language forms part of a larger system for the communication of wants and desires, attitudes and affect, requests and needs. Language itself is a product of social interaction, and in learning a language, infants learn how to interact, initiate social exchanges, respond to others in maintaining such exchanges, and how to end exchanges. In doing this, they receive pragmatic directions, both tacit and explicit, about how to use language – which words are appropriate (conventional) for particular purposes, which expressions, and which constructions. They hear and extract the regularities within a language, for example the correlations of lexical items to constructions, of sound patterns to morphemes and words, and of prosodic contours to structural units within constructions. Children assign meanings to the forms they isolate. They build up semantic domains by adding more words, assigning meanings to unfamiliar words, and attending to pragmatic directions about use.

2   In conversation with children

Infants are born into a social world, a world of touch, sound, and affect, a world of communication. They develop and grow up as social beings, immersed in a network of relationships from the start. It is in this social setting that they are first exposed to language, to language in use. This language forms part of the daily communication around them and to them. It regulates what they do. It tells them about the world, events, actions, objects, and relations within it, and presents them with affective attitudes to people and events. In short, language is a central factor in the social life of infants. The users of language they are exposed to provide the context in which children themselves become proficient at communicating wants and desires, affect and interest, requests and instructions, questions and observations, and commentary on all the contents of everyday life.

This chapter explores the social setting in which children are exposed to language, respond to it, and begin to use it. It is in and from *interaction* that children are offered conventional ways of expressing attitudes and of saying things, along with the conventional words and expressions for what they appear to be trying to say. And it is in interaction that children take up these words, expressions, and constructions. Language can be used for talking about needs and desires, or objects and events in the world at large; for talking about how to behave, how to act, what to say in different circumstances; for talking about problem-solving, for expository argument or explanation, for giving instructions; and for pretending, teasing, joking, or telling stories. In all these uses, language always forms part of a larger system for communication. It's therefore important to keep sight of communicative purposes and goals in looking at how children become members of the speaking community and learn in their turn how to talk with the same range of skills as adults. It is in the service of communication that children learn to break up the stream of speech into smaller and smaller elements, learning to identify clauses and phrases, words and morphemes.

Each linguistic chunk or unit carries meaning. So an important part of the analysis that children must do involves working out which meanings are carried by which forms. In solving this problem, they rely heavily on general pragmatic principles that language communities exploit to make sure their communicative systems remain both effective and fairly stable over time. Conventionality captures the fact that speakers expect a particular meaning to be conveyed by a particular form within their community. Members of a language community have in common a large stock of conventions – forms they expect to be used to convey particular

meanings. The inflections to use on a verb are conventional (e.g., *-ing* or *-ed*), as are the word-choices for talking about particular entities or events (e.g., *tiger* for talking about tigers, *circus* for talking about circuses), or the forms to use in making polite requests (*May I…* or *Could you…*). Conventionality covers all the agreements that members of a language community tacitly adhere to in using their language (Clark 1993).

Contrast captures the fact that the same speakers assume that any difference in form must mark a difference in meaning. If speakers don't use a conventional form, they must mean something different from what they would have conveyed by using the conventional form (Clark 1990,1993; Croft 2000). I return to these principles and their general role in acquisition in Chapter 6.

Language in context

When we think about learning a language, our first association is often to language as represented in "the form of grammars and dictionaries, or of randomly chosen words and sentences" (Halliday 1975:20). This view of language is misleading in two respects. First, it removes language from its social setting, and, second, it depicts it as a product rather than as a part of a dynamic system for communication. Language is best viewed as part of a broader communication system that draws not only on the speaker's utterances but also on gesture, stance, facial expression, affective display, and any other factors that contribute to success-ful interactions – successful in that speakers achieve their goals in conveying their intended meanings to their addressees.

The goals of an interaction, both local and global, are critical to how that interaction is carried on and what resources the participants use in communica-ting their intended meanings. Learning a first language, under this view, is part of learning to communicate. Other functions of language, as it is used to represent knowledge of the world, for instance, are put to use within a commu-nicative framework. That language is essentially social is critical in considering the settings in which children acquire language and the kinds of language addressed to them at different stages in development.

What properties of language use and language structure distinguish adult–child conversations – and hence *child-directed speech*[1] – from conversations between adults? Child-directed speech presents a major issue in research on acquisition because of the theoretical claims that have been made about its role in acquisition. While all researchers agree that children need to be exposed to language to start in

---

[1] I use the term child-directed speech in preference to other terms that have been used in research on this topic, including "motherese" (mothers are not the only people who talk to babies and young children) and "parentese" (other people also talk to young children). Both of these share an unfortunate echo of words like "bureaucratese" and "journalese." The term "input" lacks the sense of language used in communicative exchanges and any notion of cooperative exchange. And "infant-directed speech" is too limited in scope since the claims made here are not restricted to infants.

on acquisition, there is much less agreement on the form that this exposure must take. The nature of child-directed speech itself has been a matter of debate. For many years, Chomsky and some of his colleagues assumed that adult speech to young children offered at best a degenerate version of a language – such speech was full of errors, hesitations, breaks in construction, retracings, pauses, and other disfluencies, repairs to vocabulary, to pronunciation, and so on, to the extent that children would necessarily have great difficulty both in learning what might be systematic in a language and in discerning what the structures were. This view derived from a 1959 study of language production by Howard Maclay and Charles Osgood, who analyzed the transcripts of a psycholinguistics conference and extracted all the pauses, disfluencies, hesitations, and repairs in the talks and discussions. Their characterization was assumed to be representative of all adult speech. This general argument has come to be known as "poverty of the stimulus" and has been used to support the view that children must therefore be innately endowed with certain kinds of linguistic knowledge.[2] But Maclay and Osgood recorded academics speaking at a conference, not adults talking to young children.

At the same time, sociolinguistic research showed that adults are attentive to their addressees and use different styles or registers accordingly. In general, speakers have control over a variety of different ways of talking – the way they talk to babies, to foreigners, to pets, and so on – and this varies with the addressee, the occasion, and even the topic under discussion. One question here, then, is whether adult-to-adult interchanges at a conference are comparable to exchanges between adults and infants, adults and two-year-olds, or even adults and five-year-olds. Mightn't their language be modified by the nature of their addressees, in particular their age and expertise, and even by the topics talked about? Since the way adults talk to each other depends on how well they know their addressees, their relationship to them, their relative ages, the social setting, and just what they are asking them to do, mightn't this hold just as strongly for adult speech to infants and young children? By looking at just how adults *do* speak to children, one can better assess the force of Chomsky's position versus the sociolinguistic position. It turns out that child-directed speech is often singularly well tailored to its addressees, highly grammatical in form, and virtually free of errors. This makes for a rather different picture of its role in acquisition and the extent to which it presents a plausible source from which to learn a first language.

Even if the language addressed to young children is tailored to their level of skill as speakers, is such tailoring necessary for them to learn a first language? Could they acquire it instead from simply overhearing utterances addressed to others? Could they learn a first language from listening to the radio or watching television? Or do they need to hear language in interactive exchanges? What is sufficient versus necessary exposure for the process of acquisition? (Even if adults do modify their speech in talking to less-skilled speakers, this in itself doesn't tell us whether such modifications are needed for acquisition.) The nature of the exposure, it turns out, is

---

[2] I return to these issues more directly in Part IV.

important for acquisition. Children appear to need exposure to language in inter-active contexts. Merely overhearing does not appear to promote acquisition.

When children make mistakes during the course of acquisition, do they need explicit corrections (feedback) to learn the appropriate forms for what they had intended to say? Here too researchers have taken different positions. Some have assumed that children receive no corrective feedback, so the fact that they do eventually learn the adult versions of things must be evidence for the innateness of (some aspects of) language. Others have argued that feedback can take a number of forms and pointed out that adults often restate what children say, thereby offering conventional forms for the intended meanings and that such indirect corrective information is just as important as explicit rejection of an error combined with a corrected rendition. If children can learn from indirect correction, there should be less need to appeal to innateness here. In short, claims about child-directed speech have theoretical implications for claims about both innateness and learning in language.

Universal modifications?

This chapter begins by looking at what conversations with children are like, then takes up their most prominent structural characteristics and the prag-matic factors that motivate adult choices when they talk to young children. That child-directed speech differs systematically from adult-directed speech raises the question of whether the modifications adults make might be universal in those societies where adults talk to infants and young children. By modifying their language use, are adult speakers offering mini-language lessons? If so, are such lessons either necessary or sufficient? Are the effects of particular aspects of child-directed speech discernible in the patterns or rates of children's language acquisition? Or are adults simply concerned to make themselves understood as well as possible and to make their child-addressees understood too? What follows, for the process of acquisition, from this communicative goal? Do adult modifica-tions change with the age of the child-addressee? And under what circumstances do adults stop using them? These are some of the main questions that have been addressed in studies of child-directed speech.

Holding a conversation

Participants in a conversation need to observe a number of general conditions if communication is to be effective:

- Speaker and addressee must share a joint focus of attention during the con-versational exchange and take account of common ground.
- Speakers must take account of what their addressees know and tailor their utterances accordingly.

- Speakers must choose speech acts that are appropriate for the meanings they intend to convey.
- Participants in a conversation must listen to what others say so they can each make appropriate, relevant contributions when they take a turn.

The first condition depends on joint attention, with both speaker and addressee attending to the same focus, whether an object or event, and each aware that the other is also attending (Moore & Dunham 1995). This joint attention enables subsequent communication by allowing for coordination between speaker and addressee. It also identifies some common ground both for a starting point in the subsequent exchange and for coordinating as each participant adds to that common ground with each contribution. In conversations with young children, adults anchor their conversational contributions to objects or events physically present on each occasion. That is, they rely on physical and conversational co-presence as they add to the common ground in the conversational exchange.

The second condition requires that speakers tailor their contributions to their addressees, taking into account what they know – and this will include what they know about communicating, with or without language – and designing their utterances so they will be understood. For the third condition, that speakers choose the appropriate speech acts for the meanings they wish to convey, they need to use the appropriate forms, for that community, when they wish to refer, request, assert, promise, and so on. Finally, the fourth condition requires the speaker to make sure the other has understood, and the addressee to listen and signal understanding, as the exchange proceeds. This allows the participants to ground each utterance (add it to the common ground on this occasion) and so further both local and general goals in a conversation (H. Clark 1996; Grice 1989).

This pragmatic management of coordination in conversation pervades exchanges with young children just as it does those with adults, and it provides the general framework for acquisition of a first language. Conversational exchanges between an infant and an adult may be minimal at first, in the sense that the adult participant may effectively supply all the turns, as in the "exchange" between three-month-old Ann and her mother (Snow 1977:12) in (1):

(1)   Mother and Ann (aged three months)
        ANN: (smiles)
        MOTHER: oh, what a nice little smile
                yes, isn't that nice?
                there
                there's a nice little smile
        ANN: (burps)
        MOTHER: what a nice little wind as well
                yes, that's better, isn't it?
                yes
                yes
        ANN: (vocalizes)
        MOTHER: there's a nice noise

2A Daily routines in the first two years of life

(a) diaper or nappy changes in the first 24 months @ 6 per day = (365 × 6) × 2 = 4,380 (typical accompanying comments: "phew," "let me get this off you," "here we go," "now you're clean," "up with the feet," "lie still," etc.)

(b) naps and bedtimes in the first 24 months @ 2 per day = 365 × 4 = 1,460 (along with: "beddy-bye," "night-night," "in you go," "down you lie," "sleep tight," "tucking you in," etc.)

(c) mealtimes in the second year @ 3 per day = 365 × 3 = 1,095 (along with: "here's your bib," "upsy-daisy," "now get down," "do you want to get up?," "in you go," "another spoon," "here's the spoon," "one more," "let's wipe your mouth," "here's your cup," etc.)

(d) routine games and books with accompanying rhymes or routine utterances, several times a day in the first 24 months @ 5 per day = (365 × 5) × 2 = 3,650 (along with: "look," "here you are," "eensy-weensy spider," "peek-a-boo," "shall I tickle you?," "show me your nose," etc.)

Based on Ferrier 1978

As infants get older, parents raise the criterion for what counts as a contribution from their infants. At seven months, for example, this mother expected vocalizations and consonantal babble for Ann's turns and only continued talking herself after hearing such a contribution. By the time Ann was eighteen months old, her mother expected words (Snow 1977:18), as in (2):

(2)   Mother and Ann (aged 1;6)
     ANN: (blowing noises)
     MOTHER: that's a bit rude
     ANN: *mouth*
     MOTHER: mouth, that's right
     ANN: *face*
     MOTHER: face, yes, mouth is in your face
               what else have you got in your face?
     ANN: *face* (closing eyes)
     MOTHER: you're making a face, aren't you?

At each age, the mother treats the infant's contributions as if they initiated an exchange and then responds to them accordingly.

The range of topics in such exchanges tends to be rather small, so these exchanges have a highly repetitive flavor, not only when adults comment on repeated enactions of daily routines but also when infants themselves begin to contribute with more explicit content. The daily routines during the first two years of life and the stereotypical adult verbal routines that accompany them are both highly repetitive (Ferrier 1978) and very frequent, as estimated in Box 2A.

The point is, adults (or, in many societies, older siblings) talk to babies, infants, and young children as they look after them, wash them, feed them, play with them, and carry them around. Much of the speech addressed to these babies consists

of short, routine, repetitive utterances produced with great consistency and frequency in the same contexts, day after day. As babies get older and become able to do more on their own, these adult–child or sibling–child conversations encompass a growing range of topics, an ever more extensive vocabulary, and so a greater range of language uses.

Adults use language not only for talking about everyday activities and routines but also for regulating all kinds of behavior. They specify what children should say, how, and when across a range of social situations, from eating a meal at someone else's house to talking to a neighbor ("Say please"), dealing with a child who's taken away a toy ("You need to give it back"), or greeting a visitor or a relative; from thanking someone for a present to playing a game, reading a book ("Can you say *raisin*?"), petting an animal, taking turns on a swing ("You must take turns"), teasing, telling a joke, setting the table, or getting dressed ("Now your shoe"). Regulatory uses of language cover virtually every aspect of becoming socialized, of learning how to behave (see Berko Gleason, Perlman, & Greif 1984; Deffebach & Adamson 1994; Halle & Shatz 1994; Flynn & Masur 2007). Language is a primary vehicle for teaching children how to become members of a society.

Conversations become more elaborate as children understand more and take account of more uses of language. As this happens, children's turns come to contain more content, though the topics they raise may remain fairly limited for the first year or more of talking. These interactions are conversations, and they therefore place a special onus on the adult as the more skilled speaker. For instance, adult interlocutors have to monitor infant addressees with more care than they would six-year-olds or other adults to make sure the infants are attending to what is being talked about.

Joint attention comes first

In a successful conversation, the two participants must agree on what is being talked about. One way to ensure this is to start with the same locus of attention. But how does one make a one- or two-year-old systematically attend to what one is saying? One solution is for the adult to monitor what the infant *is* attending to and then talk about that (or use that as a starting point for talk) (e.g., Colas 1999; Gogate, Bahrick, & Watson 2000; Schmidt 1996). Alternatively, the adult can *attract* the infant's attention to something, with verbal attention-getters ("Hey!", "Look!") and gaze (Estigarribia & Clark 2007). Indeed, by age one, infants have become quite good themselves at checking on the adult's gaze, stance, and physical orientation, and are as likely to track the adult's locus of attention as adults are to track theirs (Moore & Dunham 1995).

Adults rely first on perceptual information to establish joint attention. If speaker and addressee are attending to the same object or event (say, a toy train), they can both more readily assume that their shared focus of attention is what the speaker is talking about, as both will have the train in mind. In adult conversations, addressees

check on what speakers are attending to and coordinate with them to achieve joint attention. But in conversations with young children, adult speakers often monitor what the children are attending to in order to achieve the necessary coordination (Barresi & Moore 1993; Butterworth & Jarrett 1991; Collis 1977; Murphy & Messer 1977; Stern 1977, 1985; Tomasello 1995; Trevarthen 1977). They rely on several perceptual cues in trying to establish joint attention with an infant or young child. They can follow the child's direction of gaze, so both adult and child can then see that the other is looking at the same thing; they can follow the child's pointing, so both adult and child can see that the other is also looking at the object being pointed at; and they can follow the child's body orientation towards something, so, again, both adult and child can see that the other is attending to the same thing (Clark 1997). And even young infants can track adult attention (Hood, Willen, & Driver 1998; Muir & Hains 1999).

Infants also become adept at actively soliciting adult attention. As young as six months of age, they co-opt adults as instruments to satisfy goals (Mosier & Rogoff 1994). By twelve months of age, they can get adults to open things, offer things that are out of the child's own reach, and attain a variety of goals they couldn't achieve on their own. In doing this, they first attract the adult's attention, then communicate what they want with combinations of gestures, vocalizations, and eventually words (Bates 1976; Carter 1978). In addition, as children get older, they attend more to adult intentions: monitoring of adult action and gaze emerges around twelve to eighteen months, along with explicit attempts to shift adult attention to what the infant wants (Leung & Rheingold 1981; Rheingold, Hay, & West 1976; Buresh & Woodward 2007).

Information about the speaker's locus of attention can provide essential information about the intended referent of an unfamiliar word. Baldwin (1991, 1993) presented infants under two with a new word in a situation where the infant played with one object while the adult looked at another as she named it. Unless they made use of the speaker's locus of attention, they could assign the word to the wrong referent. For instance, the adult speaker would focus on one object out of sight inside a bucket and produce an unfamiliar label ("A modi!") while the infant was attending to a different toy near at hand on the table. In these circumstances, infants, from sixteen months on, monitored the adult's locus of attention and so avoided unintended mappings for unfamiliar words. By age two, children can take account of repairs ("Uh-oh, it's not an *X*, it's a *Y*") and also distinguish intentional from unintentional actions ("Oops!") in assessing the speaker's intent (see Clark & Grossman 1998; Tomasello & Barton 1994; Tomasello & Kruger 1992).

Joint attention is supplemented by physical co-presence, the actual presence of the object or event at the locus of attention, and by conversational co-presence, the speaker's explicit reference to the target object or event. Together, these help ensure that speaker and addressee are talking about the same thing. With physical co-presence, the speaker talks about objects or events perceptually available to both speaker and addressee. With conversational co-presence, the speaker refers directly to the object or event that provides the topic of the exchange.

What evidence is there for reliance on physical and conversational co-presence in child-directed speech? First, adults rely heavily on the here and now in many of their exchanges with children. Talk about what is currently happening and about objects that are in use or in view for both adult and child helps ensure that each knows what the other is attending to and talking about. Emphasis on the here and now also limits the number of possible topics to what is physically present. This presumably makes it easier for both adult and child to track what the other is talking about.

Second, in the early stages of language acquisition, adults generally follow up on child-introduced topics rather than the reverse. A comparison of the average number of new topics introduced per hour of recording for one child, Eve, from 1;6 to 2;3, showed that her mother proposed about five new topics per hour to Eve's twenty (Moerk 1983). In effect, the child took the lead in initiating exchanges on new topics (see also Bloom, Margulis, Tinker, & Fujita 1996). And Eve's mother followed up on the topics her daughter introduced, expanding and commenting on what was already conversationally co-present.

Even very young children are persistent in their attempts to establish a new topic, trying a variety of means to get the adult to attend to the target object or event. In one exchange, Brenda (aged 1;8) produced her version of *bus* nine times in succession in an attempt to get her adult addressee to attend to a car going by outside. She had begun by saying *car* four times, which only elicited a "What?" of incomprehension; she then switched to the word *go* (twice), with no better success; and then she tried *bus*, only to have her interlocutor misidentify it as *bicycle*, which she rejected with *no* (Scollon 1976:109). But adults are often more successful, as in the exchange between Ann at 1;6 and her mother (Snow 1977:18–19) in (3):

(3) Mother and Ann (aged 1;6)
MOTHER (talking of Ann's nose):	don't know where it is.
ANN:	*Titus Titus*. [= cat]
MOTHER:	where, I can't see him.
	oh, there he is.
	yes, he's on the floor.
	Titus is …
ANN (interrupting):	*floor.*
MOTHER:	floor.
ANN:	*floor.*
MOTHER:	yes, Titus is on the floor.

Participation in a conversation requires signs that one is following what the speaker is saying, ratification of what the speaker has said, and contributions of one's own – additions to the topic at hand. This typically results in taking turns. But what counts as a turn in conversations with infants or very young children? When adults take part in a conversation, they expect speaker and addressee to alternate in making contributions and so adding new information, or in ratifying what the other has contributed (Fisher & Tokura 1995; Clark & Bernicot 2008).

Each turn is generally acknowledged by the other participant(s) in some way before the current speaker continues. Acknowledgements may take the form of an "uh-huh" or a head nod, or they may involve more extensive exchanges ("Did you mean *X* or *Y*?" or "That's which *X*?"), or even a full response to a question or request. So what happens in exchanges with babies or one-year-olds?

As we saw in the exchange between three-month-old Ann and her mother, turns are imposed on very young participants. An adult talking to a two- or three-month-old will count a burp, a smile, or a leg kick as a turn; in fact, adults typically say something, then wait for the baby to do something, and then resume talking. But as babies get older and extend their repertoire of actions, adults tend to up the ante. A four- or five-month-old must smile or kick; a six- or seven-month-old must vocalize; an eight- to ten-month-old must babble. That is, adults wait for an appropriate level of reaction before going on talking. Once infants begin to produce their first words, adults raise their expectations still further: now only a word (or perhaps a babble sequence) will do.

As infants begin to make more of a linguistic contribution to what is going on in the interaction, adults ask for more and more explicit expression of the meanings intended. Compare the two exchanges between Richard and his mother while looking at a book, the first in (4), when Richard was 1;1.1, the second in (5), some months later, when he was 1;11 (Bruner 1983:78, 86):

(4)   MOTHER: Look!
          RICHARD (touches pictures)
          MOTHER: What are those?
          RICHARD (vocalizes a babble string and smiles)
          MOTHER: Yes, there are rabbits.
          RICHARD (vocalizes, smiles, looks up at mother)
          MOTHER: (laughs) Yes, rabbit.
          RICHARD (vocalizes, smiles)
          MOTHER: Yes. (laughs)

In this exchange, Richard's linguistic contributions are minimal, yet his actions – touching the picture, vocalizing, looking at his mother, smiling – are clearly appropriate to the interaction and are treated as turns. Ten months later, Richard's turns contain identifiable content, as he and his mother actively negotiate over what to call the animals in the picture they are looking at:

(5)   MOTHER: What's that?
          RICHARD: *'ouse*.
          MOTHER: Mouse, yes. That's a mouse.
          RICHARD: *More mouse*. (pointing at another picture)
          MOTHER: No, those are squirrels. They're like mice but
                   with long tails. Sort of.
          RICHARD: *Mouse, mouse, mouse*.
          MOTHER: Yes, all right, they're mice.
          RICHARD: *Mice, mice*.

As children learn to make relevant contributions, they become more skilled at taking turns, at acknowledging the contributions of others, and at ratifying them. They learn when and how to make their contributions in relation to other speakers, and how and when to acknowledge the information offered by another. Acknowledgements can take the form of no more than an "uh-huh" or a head nod, or they might involve something more extensive ("Did you mean *X* or *Y*?" or "Which *X*?"), even a full clarification question, or the supplying of some requested information. The general notion of reciprocity and alternation seems to be established early through a variety of interchange types in "exchange games," notably games of give-and-take and peek-a-boo that emerge around nine months of age (e.g., Rheingold, Hay, & West 1976). The content of the child's turn needs to be pertinent to the topic that has been established. This is probably easier for children when they themselves have initiated the topic than when the adult has done so. Yet even two-year-olds will interrupt exchanges between their parents and older siblings with pertinent comments (Dunn & Shatz 1989). Remember that, from about age two on, more conversational exchanges are initiated by children than by adults, so children more often choose the topics that get talked about.

In summary, the give-and-take of conversation is imposed on babies and young children, as if to show them from the start how to be a partner in such exchanges. Then, as infants become able to make more of a contribution to what is going on, adults ask for more explicit expressions of the meanings intended. Children who begin with gestures and minimal vocalizations gradually approach conventional forms of expression, such as *look*, *that*, or terms for object categories, such as *dog*, as they get older (e.g., Carter 1978, 1979). In effect, children become more and more skilled as conversationalists.

How soon can we be sure that children are intent not just on achieving some goal but also on making sure their addressees have understood them? In many cases, the evidence is difficult to evaluate, and some researchers have concluded that young children have conversational goals but do not necessarily take account of what their addressees do or don't understand (Shatz 1983). Others have argued that even nonverbal infants are intent on making others understand. Golinkoff (1986), for instance, argued that infants initiating negotiations, rejecting incorrect interpretations of their nonverbal signals, and creatively repairing failed signals all suggest they are trying to make themselves understood to others (see also Marcos 1991; Marcos & Kornhaber-Le Chanu 1992). Others have argued that these actions can be explained by infants wanting to change someone's behavior rather than by their wanting to make the adult understand them (Shatz & O'Reilly 1990).

If children are intent on achieving understanding in addition to their expressed goals, they should repair misunderstandings whether or not they achieve their goals. Shwe and Markman (1997) therefore looked at the repairs and clarifications made by two-and-half-year-olds when they either did or didn't get a toy they wanted, and where they had either been understood or misunderstood. They reasoned that if children clarified their requests more when misunderstood than when understood, even if they had got the toy, this would be evidence that they

were taking account of the addressee's comprehension over and above their specific goal. Overall, the two-year-olds repeated the term for the toy they had requested more often when the experimenter expressed misunderstanding than when she expressed understanding; and they verbally rejected the toy they were given more often when she misunderstood than when she understood. As expected, they never repeated their request when the experimenter understood them and they got the toy they wanted. They repeated at a relatively low rate when she expressed understanding but didn't give them the desired object. But when she expressed misunderstanding, they persisted in clarifying what they'd wanted more often when they got the right toy than when they got the wrong one. So these children offer clear evidence that they care – not just about the goal but about communicating their intentions.

This concern for communication also leads two-year-olds to modify their requests in accord with what their mothers already know. Consider how two-year-olds asked their mothers for help in retrieving an object out of reach under two different conditions: In one, the mother had seen where the object was placed; in the other, the mother hadn't seen this because she was outside the room or had her eyes covered (O'Neill 1996). They labeled the object, labeled its location, and gestured to its location significantly more often when the mother hadn't seen where it was put than when she had seen this. This tailoring of utterances to what the mothers knew offers further evidence that two-year-olds are intent on communication when achieving their goals. If the goal alone had been paramount, they should always have offered all possible information about the target object and its location, but they didn't. They made use of what they knew the other person did or didn't know (see also Ganea & Saylor 2007).

Child-directed speech

What structural characteristics distinguish child-directed speech from adult-to-adult conversation? And, to what degree do the modifications that speakers make stem from their attempts to make themselves understood to less-skilled users of the language? If adult modifications depend on reactions and responses from their addressees, those modifications should change as child-addressees become able to provide increasingly appropriate responses and evidence that they have understood. And the modifications offered to one- and two-year-olds would presumably no longer be offered to five- or six-year-olds, since older children would be more likely to understand what is said to them.

Pitch and intonation

When adults talk to young children, in many languages they appear to favor higher pitch and to use exaggerated-sounding intonation contours. Effectively, they may double the range for intonation – in English from about three-quarters

Table 2.1 *Mean fundamental frequency of adult speech*

Addressees	Mean fundamental frequency (Hz)
Group 1: speech to 2-yr.-olds	267
Group 1: speech to adults	198
Group 2: speech to 5-yr.-olds	206
Group 2: speech to adults	203

Source: Garnica 1977:73. Used with permission from Cambridge University Press.

of an octave to one-and-a-half octaves – and produce higher intonational peaks with steeper rises and falls. This gives the effect of exaggerated intonation patterns. Acoustic investigations of infant-directed speech have shown that it typically displays higher overall pitch, wider and smoother pitch excursions in intonation contours, slower tempo, greater rhythmicity, longer pauses between utterances, and greater amplitude than adult-directed speech (e.g., Fernald *et al*. 1989; Grieser & Kuhl 1988).

Do infants pay greater attention to speech with such characteristics? The answer appears to be yes: They show a clear preference for it, from an early age, over adult-directed speech (e.g., Fernald 1985; Panneton Cooper & Aslin 1990; Werker, Pegg, & McLeod 1994; see also Zangl & Mills 2007). Are they attentive to the higher pitch? To the slower tempo? To the more extensive pitch excursions? In studies designed to find out whether infants were attending to the pitch, amplitude, or durational effects in adult-to-infant speech, Fernald and Kuhl (1987) gave infants resynthesized versions of speech to listen to. They found a preference for adult-to-infant speech only when the infants listened to the fundamental frequency "envelopes" of the two types of speech (adult-to-infant vs. adult-to-adult). Infants appear to be more attentive to very high pitch in speech, and the younger they are, the more attentive they are (Werker & McLeod 1989). But high pitch alone can't account for infants' attention since, in some languages, it does not occur in infant-directed speech (Bernstein-Ratner & Pye 1984). Other factors must also play a role here, such as the deliberate use of all sorts of communicative devices to attract and hold infant attention through facial expression, eye contact, touching, pointing, and so on (Stern *et al*. 1983; Werker, Pegg, & McLeod 1994).

Adults continue to use higher pitch with young children. In one study of English speakers, for instance, when adults were recorded talking to two-year-olds versus other adults, they used higher pitch to the two-year-olds than they did to the adults, across a range of speech activities. A second group of adults showed little difference in the fundamental frequency they used to five-year-olds versus adults (Table 2.1).

Why use higher fundamental pitch in speaking to younger children? When four-month-old infants are given the choice of listening to infant-directed speech (higher pitched) versus adult-directed speech, they show a clear preference for the infant-directed speech in that they prefer to listen to the higher-pitched utterances

Table 2.2 *Pitch ranges in adult speech*

Addressees	Narrowest range (semitones)	Widest range (semitones)
Group 1: speech to 2-yr.-olds	14	23
Group 1: speech to adults	6	13
Group 2: speech to 5-yr.-olds	11	16
Group 2: speech to adults	7	13

Source: Garnica 1977:75. Used with permission from Cambridge University Press.

(DeCasper & Fifer 1980). This suggests that sensitivity to higher pitch makes infants more attentive when they hear relatively higher voice pitch. This then allows pitch to act as an attention-getter for infants and young children. As Fernald (1989) pointed out, higher pitch may also distinguish speech directed to the infant from other background talk and noise by making that speech more audible.

The adult speakers in Garnica's (1977) study distinguished both two- and five-year-olds from adults in that they used a wider pitch range (the distance from low to high point in pitch), measured in semitones, to both groups of children than they did to other adults. Their intonation with children was more exaggerated than with adults. In fact, the narrowest range in speech to children was typically the same as the widest range in their speech to adults (Table 2.2). This perhaps is where children begin to learn what the intonational system is for their language, by learning "some of the meanings of the adult intonation system," for example, which contours signal questions and which signal assertions (Cruttenden 1994:145). Exaggerated intonation contours also attract attention by distinguishing adult speech to children from other types of conversation (and addressees).

Rate, pausing, and fluency

The steep rises and falls in intonation might also mark phrase- or clause-boundaries, along with pauses. Broen (1972), for example, analyzed the locations of pauses in mothers' speech to their two-year-olds and five-year-olds compared to conversation with other adults and found that between 75% and 83% of pauses in speech to the children occurred after terminal contours at the ends of sentences (that is, final falling intonation contours), compared to only 51% of the pauses in conversation with adults. This difference was even more striking when she looked at the sentence-boundaries followed by pauses. In talking to their two-year-olds, mothers paused at the ends of sentences 93% of the time; with their five-year-olds, they paused 76% of the time, and with adults, they paused only 29% of the time. The pauses in speech to young children are consistently longer than the analogous pauses in adult-to-adult speech (Fernald & Simon 1984). The combination of falling intonation and a pause, then, marks the ends of utterances in a highly reliable fashion in speech to young children and so provides clear information about boundaries, both for the utterances as a whole and for the final words in those utterances.

Table 2.3 *Words per minute in speech to children versus adults*

Addressee (age)	Free play	Storytelling	Conversation
2;3–3;5	69	115	–
3;10–5;10	86	128	–
Adults	–	–	132.4

Source: Broen 1972:6. From Broen, Patricia A. 1972. The verbal environment of the language-learning child. *Monograph of the American Speech & Hearing Association* 17. © American Speech-Language-Hearing Association. Reprinted by permission.

Utterance-final position is salient to infants for another reason as well. In speech directed at infants and young children, adult speakers consistently lengthen the vowels of words they wish children to attend to. Adults talking to two-year-olds typically lengthen the stressed syllables in words they want children to attend to, for instance, when solving a puzzle, but they do this much less in talking to five-year-olds or to adults (Garnica 1977). In another study that compared infant- and adult-directed speech, adult speakers (mothers) were asked to label seven objects when speaking either to the experimenter or to their infant. The target words directed at the infants were both higher pitched and had greater syllable-lengthening than the analogous adult-directed speech (Albin & Echols 1996). Stretching out words as well as raising the pitch both seem to be designed to attract the young addressee's attention. In fact, two-year-olds appear to make use of all these cues as they interpret what adults are saying (Shady & Gerken 1999).

Adults generally speak more slowly to young children than to older ones or to adults. In her detailed study of parental speech, Broen found consistent differences in the number of words per minute in the same mothers' speech to two-year-olds and five-year-olds, in both free-play and storytelling, compared to speech to adults (Table 2.3). (Some of the features of slow speech also show up in the overly careful articulation adults use in human–computer interactions when people are trying to make a computer recognize words (e.g., Oviatt *et al.* 1998).)

The slower rate to young children is achieved through pauses rather than stretched out words. That is, adults pronounce individual words at the same speed as in adult-directed speech, but they insert more pauses (at sentence- and phrase-boundaries) in their speech to younger children. Overall, adults use fewer than four words per utterance to two-year-olds compared to over eight words per utterance to adults (Phillips 1973). The shorter sentences used to two-year-olds are also simpler in structure than those used to older children or adults. For example, Sachs, Brown, & Salerno (1976) found that adults used simpler constructions in telling a story to a two-year-old than to another adult. To the child, they used only a few coordinate and subordinate clauses (introduced by *and, when, if*, or *because*) and hardly any relative clauses (e.g., *The dog that I patted ran away*), comple-ments (*He wants to climb up*), or negations (*They didn't come*).

Table 2.4 *Mean number of disfluencies per one hundred words*

Addressee (age)	Free play	Storytelling	Conversation
2;3–3;5	0.58	0.66	–
3;10–5;10	1.61	0.77	–
Adults	–	–	4.70

Source: Broen 1972:11. From Broen, Patricia A. 1972. The verbal environment of the language-learning child. *Monograph of the American Speech & Hearing Association* 17. © American Speech-Language-Hearing Association. Reprinted by permission.

Adults are also much more fluent when they talk to young children than when they talk to other adults. They produce many fewer false starts, mispronunciations, or hesitations – about one-ninth of the rate in their speech to adults (Table 2.4).

Repetitiousness

Adults also repeat themselves a lot in talking to young addressees. One reason is that they rely heavily on a small number of constructions that combine a small "sentence frame" with a noun phrase or a nominal. Some typical examples are listed in (6):

(6)　　　Construction　　　　Example
　　　　Where's NP?　　　　Where's Daddy?
　　　　Here's NP　　　　　Here's (the) kitty
　　　　Look at NP　　　　Look at (the) doggie
　　　　That's a N　　　　That's a ball
　　　　Here comes NP　　　Here comes Danny
　　　　Let's play with NP　　Let's play with the blocks

Adults use constructions like these to introduce new words and often produce them with an exaggerated intonation contour and heavy stress on the new word in final position (Broen 1972; Ferguson, Peizer, & Weeks 1973; Clark & Wong 2002). At times they make use of question–answer pairs, both spoken by the adult, as in "Where's the ball? Here's the ball." These adjacency pairs, normally produced across a pair of speakers, are quite frequent. Children soon learn their part, the kind of response needed after the adult produces the first part of such a pair. For example, they respond to *How many*-questions with a number, as in "How many frogs do you see?" – "Two" (regardless of the actual number depicted), or to *What colour*-questions with a colour term, as in "What colour is your ball?" – "Red" (even if it isn't). That is, children learn the appropriate kind of response before they have fixed the reference for terms such as *two* or *red* (Clark 2006; Clark & Nikitina 2009).

Adults also repeat themselves with small variations when they ask young children to do things, as in (7). Repetitions like this in English are three times more frequent in speech to two-year-olds than in speech to ten-year-olds (Snow 1972; see also Shatz 1978a, 1978b).

(7)   ADULT (trying to get a two-year-old to pick up some blocks): Pick up the red one.
           Find the red one. Not the green one. I want the red one. Can you find the red one?

In highly inflected languages like Turkish or Finnish, adults often rely on variation
sets, utterances with much the same semantic content and intent, as in (7), but
with extensive changes in word order from one utterance to the next. Küntay and
Slobin (1996) argued that such variations help children identify the stable elements
like verbs that recur from one utterance to the next, and so could offer important
information for identifying chunks as words (see also Bowerman 1973a).

    In summary, adults consistently produce shorter utterances to younger addressees,
pause at the ends of their utterances around 90% of the time (50% in speech to adults),
speak much more fluently, and frequently repeat whole phrases and utterances when
they talk to younger children. They also use higher than normal pitch to infants and
young children, and they exaggerate the intonation contours so that the rises and falls
are steeper over a larger range (up to one-and-a-half octaves in English). The
grammaticality, fluency, and simplicity of the language addressed to young children
shows that earlier assumptions about child-directed speech were simply wrong.

    Adults streamline their delivery when they speak to young children, and they
appear to do this more the younger the child, with the most careful delivery
directed at children just starting to speak. This streamlining may be attributable
in part to the greater ease of planning and producing really short utterances. This
would also account for the relative absence of speech errors in child-directed
speech compared to adult-to-adult speech (Broen 1972).

    While this summary captures some of the main structural characteristics that
have been observed in speech directed to young children in various Western
societies, it does not consider all the modifications adults make, nor why they
might make them when and how they do. We turn next, therefore, to some of the
main functions that adult modifications seem to serve.

Functions of child-directed speech

           Why do adults speak to infants and young children differently from
other adults? What motivates the modifications they make? I will argue here that
their modifications help speakers get and keep their addressees' attention. These
addressees are young and unskilled as speakers, and have only a limited know-
ledge of the language around them. The changes adults make in how they talk
seem designed to ease their communication with such addressees. First, they need
to make sure they and their addressees are attending to the same objects or events,
that there is joint attention on the target topic, so they can then direct it to the
relevant event. To do this, they use devices to signal that that addressee and no one
else is the intended addressee: They use a vocative (the child's name) or an
endearment (*Sweetheart!*); they use a deictic term as a summons (*Here! Look!
See!*); and they mark their utterance with higher than normal pitch, for example, to
distinguish it from utterances that might be designed for other addressees.