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 constructions. 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 communication. Or they could start from nothing and discover what is (and isn't) represented

1

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 pro-
duces 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