Chapter 9

Language Acquisition

Contents

9.1 Introduction .............................................. 175
9.2 The Earliest Stages ..................................... 176
  9.2.1 Perception ........................................... 176
  9.2.2 VOT Contrasts ...................................... 177
  9.2.3 Production ........................................... 179
9.3 Later Stages ............................................. 179
9.4 Acquiring a Lexicon .................................... 181
9.5 Theories of Acquisition ................................. 182
9.6 The Critical Period Hypothesis ....................... 184
9.7 Summary .................................................. 185
9.8 Exercises .................................................. 185

9.1 Introduction

The process of language acquisition sits at the very center of the major concerns of modern linguistic theory. It is fundamental both to the problem of characterizing the nature of grammar (some part of which is assumed to be an innate endowment of humans, another part learned) and to an understanding of language change (and therefore, of language variation).

While linguistic research has made considerable progress in the study of certain aspects of acquisition, the picture is far from complete. This is partly because children, especially young children, are rather difficult to work with, and partly because experimentation on human subjects is not always possible in general. (Oddly, parents will not
generally allow their children to undergo useful experiments which involve procedures like injecting radioactive fluid into their children’s brains so that we can get an image of the activity in there, solely in the interest of scientific curiosity!) In addition, children are even worse than adults at conveying anything useful about what’s going with them linguistically. In the early stages they can’t talk (and we are unable to extract information from bubbles and various gastric noises). Later on, they talk after a fashion, but it turns out they generally have no idea what you’re talking about when you ask them perfectly reasonable questions (like: ‘Is that an allophone of this?’). Consequently, with the exception of some nifty non-invasive experiments for speech perception, most of our data on acquisition comes from the same sources as our data for adult grammars – production and comprehension – with all of the associated drawbacks. For reasons we will examine shortly, using children’s production is even more fraught with uncertainty than using adult production and children’s comprehension data is not appreciably better.

Before discussing acquisition data and methodology in more detail, we will outline a few points about knowledge of language which are relevant to the question of which aspects of that knowledge are innate (part of UG), and which aspects must be learned from the environment. An innate component is suggested by the following:

- Aspects of linguistic knowledge appear to be present at birth.
- The development of language mirrors the development of other, innate abilities.
- The system is too complex to be acquired by children in the time that we know they acquire it.
- The data that children receive is insufficiently robust to allow them to construct a grammatical system without relying on some innate knowledge.

A learned portion is suggested by the very obvious fact that children acquire whatever language (or languages) they are exposed to in their environment. It is not quite as straightforward as simply memorizing vocabulary items, of course, but it seems to be true that a child will be unable to acquire language if he or she is not in some language environment.

9.2 The Earliest Stages

9.2.1 Perception

While infants are not known for their stimulating conversational repartee, they do demonstrate some surprisingly sophisticated, and language-specific, apparently, knowledge of speech sounds. To study early (from within a few days or even hours of birth to about 4 months) infant speech perception, a particular type of experimental procedure was developed – the High Amplitude Sucking test (HAS). This test takes advantage of one of the
few activities babies of this age are capable of, as well as of the intriguing fact that babies appear to suck faster on pacifiers when they’re interested in or excited by something than when they are bored. In the HAS experiment, a pacifier is attached to a pressure transducer which records sucking rate. A baseline sucking rate is established for each infant and subsequently the infant is exposed to some stimulus. When first exposed to the new stimulus, the sucking rate increases, but as the same stimulus is played repeatedly, the sucking rate drops off (the infant becomes habituated). When the ‘old’ stimulus is replaced with a new stimulus, the sucking rate increases (and then drops off again after a certain period of time).

Using this technique, experimenters have tested infant perception of a number of differences between natural language speech sounds. For example, when infants are played the sound [ba] in a HAS experiment and then the stimulus is changed to [pa], infants react with increased sucking rate, indicating that they perceive [pa] to be a new stimulus (thus they distinguish between [ba] and [pa]). This same type of experiment has been performed on infants from a range of different language environments and the stimuli have included a range of natural language speech sounds. The results of these experiments suggest that infants from the earliest ages are already sensitive to the possible sound distinctions in language. Moreover, their discrimination, at this early age, is not at all dependent upon the language environment that they have been in. For example, English environment infants are as capable as Hindi environment infants at distinguishing between sounds that are present in Hindi but not in English and vice versa.

Results such as these suggest that the ability to perceive distinctions in natural language speech sounds is innate rather than learned. On the other hand, one might perfectly well argue that all that these experiments show is that infants have very acute hearing and that they are not discriminating between speech sounds, merely between sounds. In order to try and resolve this question, experimenters have run additional tests based on what we know of adult perception of certain distinctions.

9.2.2 VOT Contrasts

In the HAS experiment described above with [ba]/[pa], the difference that infants are perceiving is one of voicing on the initial consonant. The true description of what ‘voicing’ consists of for stops is a little more complicated. It turns out that the real cue for the listener is not simply whether or not there is vocal fold vibration but, instead, exactly when that vibration starts with respect to when you release your lips (in the case of a bilabial stop). For a listener to perceive a bilabial stop as voiced (i.e. as [ba]), vocal fold vibration must start anywhere between about 10 milliseconds before (roughly at the start of the bilabial closure for [b]) to 10 milliseconds after the lips are released. In the case of [pa], speakers usually wait until about 20 milliseconds into the [a] before they start voicing (that is, 20 milliseconds after the release). What we have been describing here is Voice Onset Time (VOT). VOT is literally when voicing starts with respect to the release of the stop closure (in the case of bilabials, that is when you open your lips). Again, for bilabials,
if you start voicing anytime between 10 milliseconds before the stop release (-10 msec. VOT) and 10 milliseconds after (+ 10 msec. VOT), an adult listener will hear [ba], while if you start voicing sometime after 20 milliseconds (+ 20 msec. VOT) after the stop release, adults will hear [pa]. Experimenters based a number of acquisition experiments upon these observations of adult behavior, as well as on one additional factor which we will talk about at the same time as the experiment itself. These experiments were geared toward answering the question that arose from the perception discrimination experiments described earlier, i.e., is infant reaction simply accurate audition or is it the result of knowledge particular to the language faculty.

The experiment proceeds in a similar fashion as the simple [pa]/[ba] experiment but with some fine tuning. First, a series of [ba]'s is played to the infant but the VOT of the [ba]’s is different. The first set contains [ba]’s which start voicing at 10 milliseconds prior to release. When the infant’s sucking rate decreases, a second set of [ba]’s which start voicing 10 milliseconds after release is played. Critically, the infants show no interest in this 20 millisecond change of voicing – they do not increase sucking rate when the second type is introduced. Similarly if you play a series of [pa]’s, the first set of which has voicing starting 20 milliseconds after the stop release and the second set with voicing starting 40 milliseconds after the stop release, the infant sucking rate continues to fall as if there was no change in stimulus. Note that, in both cases ([ba]’s and [pa]’s), there was a 20 millisecond difference in VOT between the first and second sets.

One more set of stimuli, which differ only by 20 milliseconds, is now added to the picture. The infant is played a [ba] with voicing which starts simultaneously with release. Then a second stimulus is introduced in which voicing starts 20 milliseconds after release. In this case, the infant does react as if there is a new stimulus and the sucking rate increases. Now in this and the previous two cases, voicing was shifted by exactly 20 milliseconds but the the infant treated two of the 20 millisecond shifts as if nothing new happened, unlike the third shift. Since the stimuli were equidistant from one another, the only explanation for the fact that the child treats only one pair of them as if they were different is if the child is particularly sensitive to the difference between 0 millisecond delay (after stop release) and a 20 millisecond delay (after stop release) – precisely the difference between a [b] and a [p] in adult language. This phenomenon is called ‘categorical perception’ and, unlike the first results cited for simple [pa]/[ba], cannot be explained by simply saying that the stimuli were acoustically different. (Note, again, that every shift was between acoustically different tokens of 20 msec. but only 1 of these shifts was treated by the infants as making a difference.) These results seem to indicate something about the presence of knowledge specific to language in infants. To emphasize this, HAS experiments have been done using pure tones, rather than speech sounds. No categorical difference of the type described above exists for the tone-based experiments.1

---

1Nothing is ever clear-cut, unfortunately. Chinchillas are also very good at categorical perception experiments, whatever that fact means.
9.2.3 Production

From the production perspective, there is little going on at the earliest ages. Most of the sounds produced by infants up to about 3-5 months fall into the attractively labelled ‘vegetative noises’ category. After this time, however, infants begin to make sounds which are somewhat more related to speech sounds. At this point, their production is called ‘babbling.’ The babbling sounds are not linguistically significant in that they are not produced consistently with respect to naming anything, however listeners hear these sounds as speech-like. While babies in this stage produce a very wide variety of speech-like sounds, the majority of the consonant sounds they make are from the set \{h d b m t g w n k\}. Once again, this set is language environment independent – all babies babble similarly regardless of the sounds heard in the particular environment. Given which sounds are most ‘popular’, as indicated by the set above, it seems likely that babies are producing most often the sounds which require the least sophisticated motor coordination. Sounds which require fine motor control and airflow regulation (many fricatives, for example) are typically only produced in an adult-like fashion much later (possibly much much later). In general, the babbling stage is felt to be a type of ‘practice’ for the real thing later on.

9.3 Later Stages

From about 12 to 18 months, children start producing some recognizable speech (all things considered). Their earliest utterances consist of single words. As a result, this is called the ‘one word stage.’ The assumption is that children actually intend something like a sentence but can only manage a little part of it. The latter is due, in our opinion, to general limitations that it is well known children have, such as limited short-term memory, limited motor coordination, limited attention spans, and so on. If one is limited to one word, a good strategy is to pick the word that will be key in getting your message across – this appears to be what children do.

Interestingly, comprehension far exceeds production at this point in time. In general, production does not catch up to comprehension for a number of years although the gap becomes narrower as time wears on. (One could conceivably make the point that production never actually catches up to comprehension, of course, given adult performance vs. competence.) Children at the one-word stage are able to interpret at least simple sentences relatively well even though they cannot produce such long utterances themselves. In a frequently depicted experiment, English environment children at the one-word stage are placed in front of two TV screens. One of the screens will show Big Bird tickling Cookie Monster and the other will show Cookie Monster tickling Big Bird. (Success is predicated upon familiarity with these characters, obviously.) A loudspeaker will announce something like ‘Look! Big Bird is tickling Cookie Monster!’ and the children turn toward the appropriate screen and concentrate on it for some time. This shows that the children

\(^2\)Note that an awful lot must have gone on to get to this point from the earlier stage. The details of all of this (those that we’ve managed to figure out, anyway) are too many to include here.
already have some relatively advanced knowledge of the syntactic structure of English where the subject (the one doing the action) comes first in a simple sentence. This knowledge is certainly not displayed in their very limited production at the time, supporting the position that production should not form the basis for determining children’s language development (a development of the mind, not the articulators).

A few months after the one word stage, the next stage in production occurs. This is the ‘two word stage’ characterized by – you guessed it – utterances of only two words. In this case, the strings are frequently, although not necessarily, composed of a noun and a verb.

- Allgone sock.
- Baby chair.
- Doggie bark.
- No eat.
- Throw ball.

Again, these two word strings are assumed to represent more complete utterances (sentences) but due to the limitations mentioned above, only a couple of key words are there.

There is no three word stage to follow the two word stage. Instead, some months later, utterances which have varying numbers of words but virtually no function words begin to be produced. Because this is similar to the style of old-time telegraph messages (where all but the most important words were omitted for reasons of cost), this stage is called the ‘telegraphic stage.’ This stage might include strings like the following.

- Man ride bus today.
- Me wanna show Mommy.
- Me put it back.
- No do that again.

Beyond this point, no very clear distinctions can be made in terms of stages of production. The number of vocabulary items soars and there is constant progress toward approximating adult speech. By the time children are about 5 years old, their grammars appear little different from adult grammars.3

3Note that we said ‘grammars’, not ‘articulation’, ‘vocabulary’ or ‘rhetorical ability.’
9.4 Acquiring a Lexicon

Building up a lexicon is, in some ways, not only the most fundamental but also the most difficult part of acquiring language. For necessary reasons, we believe that UG provides much of the structure of the computational system, so that it doesn’t have to be learned. However, the lexicon must just as obviously be acquired through experience. Even if one assumes (as we do) that there is an innate set of semantic features available in UG, the tricky question is: How do children (and adults) form hypotheses about which semantic features should be grouped together to make up a coherent semantic representation for a lexical entry? There is evidence from production and comprehension that children in the acquisition stage do not ‘zero in’ on an adult-like representations the first time around. Interestingly, children seem both to overgeneralize and undergeneralize. An example of the former is a child who refers to any number of (non-dog) animals as ‘dog.’ An example of the latter is a child who only refers to his or her own dog as ‘dog.’ There are a number of possible explanations for both of these phenomena, only some of which have anything to do with the accuracy of children’s semantic representations. (We will leave these to you as a ‘thought exercise’ for now and return to it in question form at the end of the chapter.) In spite of these apparent false starts, the vast majority of children converge on something like the adult semantic representation of ‘dog’ as far as we know. In addition, however, they all perform what seem like magical leaps in the area of semantic representation. When they finally converge on ‘dog,’ they are able not only to identify virtually all and only the disparate members of the canine species as ‘dogs’ but also to identify stuffed animals that look like dogs as ‘dog,’ photographs, drawings and paintings of dogs as ‘dog,’ and things that bark in the night as ‘dog.’ Notice that the physical reality of stuffed toys, pictures, and noises are quite different both from one another and from a live dog. The evidence of ‘doghood’ must come from a sufficient number of abstract semantic features that one can remove a good many features and still be left with something that is ‘dog.’

This very brief discussion of semantic features brings up the fundamental question of whether or not any two individuals have exactly the same featural representation for a particular morpheme in the same language, let alone for what we think of as an equivalent morpheme in some other language (cf. ‘dog’ and the French equivalent ‘chien’). The assumption that there is a universal semantic feature set available does not entail the further assumption that everyone constructs identical semantic feature bundles. It seems likely, once again, that functional morphemes and content morphemes may be different in the domain of acquisition. Functional morphemes tend to be much more circumscribed, something that would make them easier to acquire. For content morphemes, on the other hand, experience must play a large role in what features are chosen to represent a particular morpheme. Since no two individuals have identical experiences, one might hypothesize that there exist a number of differences between semantic feature bundles across individuals. This is probably more likely in the case of abstract notions, such as ‘like,’ ‘despair,’ ‘acquaintance,’ and ‘friend’ than for concrete objects such as ‘dog’ or ‘table,’ but may well be true of all content morphemes. However, if such differences exist,
they seem to be, for all practical purposes such as casual conversation, insignificant.

### 9.5 Theories of Acquisition

Over the years, a number of theories have been put forth to account for the fact that children acquire language. One of these theories was the imitation theory. It held that children learned language by imitating the people speaking around them. This theory has an obvious disadvantage in that it seems to predict that we should only be able to say things that we’ve heard others say (if all our knowledge is due to imitation, what else could we do?). As the examples from the two word and telegraphic stages suggest, children are doing something other than imitating people around them. Few adults produce sentences such as ‘Man ride bus today’ or ‘Allgone sock.’ Even the more ‘subtle’ aspects of language development cannot be due to imitation. For example, children’s morphological development shows that they at some point deduce the existence of regular, productive morphology and incorporate the appropriate rules into their grammar. This leads to output such as the following.

<table>
<thead>
<tr>
<th>Earlier stage</th>
<th>Later stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>dogs</td>
<td>dogs</td>
</tr>
<tr>
<td>cats</td>
<td>cats</td>
</tr>
<tr>
<td>mans</td>
<td>men</td>
</tr>
<tr>
<td>childs</td>
<td>children</td>
</tr>
<tr>
<td>kicked</td>
<td>kicked</td>
</tr>
<tr>
<td>goed</td>
<td>went</td>
</tr>
<tr>
<td>runned</td>
<td>ran</td>
</tr>
</tbody>
</table>

It is only at a later stage that lexical exceptions are ‘recorded’ (stored as exceptions in the lexicon) and then produced appropriately. There is no explanation for non-adult forms in the first column under a theory which says that children imitate those around them (i.e. adults).

Another theory, the reinforcement theory, popular with behavioral psychologists earlier in this century, held that children made some noise, and if they received positive reinforcement from their caregivers for making this sound, they made it some more. Eventually, through the process of continual adult reinforcement, children came to say words and eventually sentences. An interesting dialogue between parent and small child reveals some problems with this theory.

Child: Nobody don’t like me.

Parent: You mean ‘nobody likes me.’

Child: Nobody don’t like me.
9.5. Theories of Acquisition

Parent: Say ‘nobody likes me.’
Child: Nobody don’t like me.
Parent: Say ‘nobody.’
Child: Nobody.
Parent: Say ‘likes.’
Child: Likes.
Parent: Say ‘me.’
Child: Me. Nobody don’t like me.

Note first that any normal (i.e., non-linguist) parent would respond to this child’s pathetic assertion of lack of self-worth with a supportive ‘Oh, come on, everybody likes you.’ or some such statement. This dialogue thus reveals two difficulties for the ‘reinforcement theory.’ First, parents generally reinforce children not for the form of their assertions but for the content. For example, when, as a child, one of the authors said ‘Toni is stupid’ (Toni is his sister), he did not get a nice pat on the head from his mother because of his proper use of the verb ‘to be.’ He got, instead, some negative reinforcement in the form of a swat on the behind. On the other hand, when he said ‘I wuves you mommy’ he got a big hug, in spite of the ill-formedness of his utterance (and the fact that he had his eye on some more candy). Reinforcement will never explain the acquisition of the grammar.

This dialogue is equally bad for the ‘imitation theory.’ The parent offers several, carefully presented models for the child to imitate and the child stubbornly persists in saying it his way as opposed to imitating. And again, who could the child be imitating when he says ‘nobody don’t likes me’? This particular child was born into an environment in which there were no speakers of any dialect which produced such forms. Instead, what such dialogues seem to reveal is the same aspect of acquisition that the morphology examples revealed. At any particular point, the child has a grammar – it’s just a different grammar than an adult grammar. The child’s motivation is neither to get some particular reinforcement nor to accurately imitate an adult, but rather to express him/herself, using the grammar he has constructed.

Regardless of what kind of input they receive, a child learning an English type grammar will produce sentences like ‘nobody don’t like me.’ This can be explained if this type of output reflects a necessary intermediate step between the innate linguistic system, present in all speakers at birth, and the fully-acquired English-type grammar. The innateness hypothesis, as this theory is called, has many distinct advantages over its earlier competitors. Theories which have the child starting out with nothing in their heads (so-called tabula rasa or ‘clean slate’ theories) not only can’t explain the HAS results on VOT, they would also never predict dialogues such as the one cited above. The fact that
sentences with this structure are uttered by *every* acquirer of English tells us something important: there is an innate linguistic system. The role of caregivers and others within hearing distance of the child is far less obvious than many people supposed. Speakers in the child’s environment (caregivers, siblings, and so on) certainly do provide the child with all important data, but they cannot influence what the child *does* with the data they receive – this appears to be regulated by some unconscious, internal acquisition mechanism of the child.

If we take the facts above together with the obvious complexity of human grammars (as you’ve seen in this course) and the relatively poor quality of the data that children have to work with (consider all of the performance errors in natural speech), a clear conclusion may be drawn. Children are really good at learning languages because they are preprogrammed to do so. They are born with innate knowledge of precisely what kinds of things they might find in a grammar and their learning task is more a matter of weeding out elements they don’t need (because it isn’t in the particular language they’re being exposed to) than a matter of discovering facts of great complexity on the basis of really bad evidence. While the individual lexical items of the language to which they are exposed must be learned, the structural possibilities for all human languages are present in their minds at birth – which we term UG.

It seems appropriate to end this section by referring you back to the waveform and spectrogram at the beginning of Chapter 2. Those show the physical reality of language – no discrete sounds, no words, no sentences, no meaning – just air pressure variation. The human acquirer gets *air pressure variations* and ends up with a mental system of incredible complexity before they can successfully tie their shoes. An innate component to knowledge of language seems a logical necessity given this.

### 9.6 The Critical Period Hypothesis

The developmental stages described in this chapter are stages that every human child goes through (barring severe and specific neurological abnormalities and assuming normal exposure to language data in their environment). While none of the ages are absolute for any one stage, the order of acquisition is essentially the same for all children and the ages at which certain aspects of language develop are approximately the same. In this sense, language acquisition is similar to walking. All children learn to walk in pretty much the same way at pretty much the same time – speeding up the process through outside intervention is not possible, as far as we know. Many innate ‘abilities’ are on-line, in some sense, at birth. Human hearing is an example. Others develop into their adult-like state soon after birth, such as vision. Still others, such as walking, arguably, take a rather longer time to become adult-like. When developmental time after birth is required, it is

---

4The pendulum has swung back and forth on whether caregiver talk (‘baby talk’, ‘motherese’) has positive or negative effects on acquisition. The tremendous variation across cultures in this domain, however, suggests that any possible effects would be insignificant.

5It is interesting to speculate on whether a child will walk if they are not surrounded by ‘walkers.’
also sometimes the case that environmental input is crucial to proper development. Vision, in at least some species, is like this. In such cases, there is said to be a ‘critical period’ for development. This brings up the interesting question of whether there is a ‘critical period’ for acquiring language. That is, does language, specifically, one’s first language, have to be acquired within a certain time frame?

There is no humane experiment one can perform in order to help answer this question. The only information that might be relevant comes from a handful of tragic cases in which children have been deprived of linguistic stimulus during the normal developmental years. The most famous and most carefully studied of these cases was of a girl, Genie, who was deliberately deprived of almost all normal human interaction through her first fourteen years. A linguist, Susan Curtiss, worked with Genie for some period of time after she was rescued from her home environment. Although Genie made considerable progress, her acquisition was ‘spotty’ and much more successful in the realm of vocabulary than of structure. Her linguistic development was unlike that of a child from a normal environment and remained far from adult-like. Few conclusions can be drawn about the notion of a critical age for acquisition, though, due to the existence of many other variables such as the impact of psychological trauma and neurological problems. In general, the multitude of factors involved in cases such as these do not allow us to get a clear picture of what might be going on.

9.7 Summary

Human knowledge of language is a system of great complexity and we are far from understanding all of its intricate mechanisms. On the other hand, how hard can it be if a three year old can do it? (Note again that, for a three year old, tying one’s shoe is an insurmountable obstacle.) The answer appears to be ‘Not hard, if you’re human.’ but ‘Not likely, if you’re not.’ Humans also seem to treat bipedal locomotion as a trivial activity whereas an attempt to fly by flapping one’s arms is doomed to failure. All species do what they are genetically programmed to do. And all species have some mechanism that allows them to communicate amongst themselves. The human linguistic system is a property of the human genetic code. Language, in the sense it has in linguistics – a mental system made up of both genetic and experiential components – is thus a defining property of what it means to be human.

9.8 Exercises

1. If a child in the early stages of acquisition refers only to his/her dog as ‘dog’, what might this indicate?

---

6The ‘critical’ part may be the maturation of some aspect of the organism, some input from the environment, or both.
a) The child’s grammar takes ‘Spot’ and converts it to ‘dog’ by regular, phonological rules.

b) The child has taken ‘dog’ to refer to the very specific set of attributes of his/her particular dog (i.e., hair of a particular length and color, bulgy brown eyes, wart on the chin, and so on).

c) The child will never converge on some adult-like meaning of ‘dog.’

d) The child initially assumes that the sound [d] refers to hairy objects, and [g] refers to 4 legged objects.

2. Which of the following is true?

a) Some morphemes have meaning and some do not.

b) Some morphemes have ‘content’ semantics and some ‘functional’ semantics.

c) All morphemes have both ‘content’ semantics and ‘functional’ semantics.

d) The semantics of simple morphemes are derived during the course of a derivation.

3. If a child in the early stages of acquisition refers to a bunch of ‘non-dog’ animals as ‘dog,’ this indicates that:

a) the child’s knowledge of language includes only phonological features at this point, not semantic features.

b) the child has heard everyone else refer to those other animals as ‘dogs.’

c) the child cannot distinguish between ‘non-dog’ animals and dogs.

d) the child has overgeneralized — later he/she will refine their semantic feature set for ‘dog’ to be more specific.

4. During the course of language acquisition

a) production is constantly ahead of comprehension

b) comprehension and production proceed at the same pace

c) comprehension outstrips production by far at the earlier stages

d) comprehension starts to lag behind production after awhile