Psych 229:
Language Acquisition

Lecture 3

Acquisition & Levels of Representation

Stages of acquisition

Sometime in the first couple of months, babies develop a kind of vocalization usually called "cooing": "goo" or "gmp" sorts of sounds, or quiet whooping. This gradually gives way at about six months to a stage called "babbling," in which the baby makes a large range of meaningless sounds, often forming strings of syllables. Frequently, babbling children even make sounds that aren't present in the language of the environment.

The consensus on babbling is that it is basically a stage in which the baby is playing with its vocal tract, with no particular linguistic intention. Even deaf babies are observed to babble; on the other hand, not all babies do it (my older daughter didn't, to my deep disappointment). Still, there are hints of proto-linguistic behavior. Babies often babble in response to being spoken to, suggesting that they are catching on to the idea of taking turns speaking in conversation. And a couple of months into babbling, the strings of sounds begin to be uttered with intonation patterns characteristic of speaking, so that the baby almost seems to be talking.

(Incidentally, deaf children exposed to sign start "babbling" with their hands, in a way very much parallel to spoken babbling, experimenting with handshape and movement.)

Sometime between ten and twenty months (with girls tending to be on the earlier side, boys later), babies really start to talk, albeit in single-word utterances.

The child's vocabulary may grow to fifty or seventy-five or a hundred words over a period of six months or so. You can list the words your child knows, and each new word is a milestone: "Hey, Beth said 'turtle' today!" Despite the limitations of this one-word stage, a surprising amount gets communicated this way.

Stages of acquisition

After a few months of this kind of talk, perhaps at two years of age or a little before, children start to put together two-word utterances, things like "Mommy sock," "drink soup," "no eat." Even though there is nothing like an adult grammar yet, we see fairly consistent use of word order, in a sort of stripped-down version of adult order. For instance, a child at this stage won't say "Mommy throw ball," because it's too long. But we may well hear the more reduced versions "Mommy throw" and "throw ball"; while the opposite orders, "throw Mommy" and "ball throw," are unlikely.

Around the same time, all of a sudden the child's vocabulary takes off. The parents can't keep track anymore of the words their child knows. The standard estimate is that a five-year-old knows on the order of 10,000 words. This means that between the ages of two and five (three years, about a thousand days), the child has averaged ten new words a day, or close to one every waking hour! Since a word may take a period of time to master, this also means the child is probably working on dozens of words at a time.

After maybe another few months of two-word utterances, we begin to see a steady growth of grammatical complexity along with vocabulary growth. The child starts constructing gradually longer and more complex sentences, and function words and inflections begin appearing. By age five the child is speaking with a very good approximation to adult grammar, though there are numerous wrinkles to be ironed out and complexities to be added by age ten or so. (And vocabulary learning continues throughout life, though at a less frenetic pace.)

Stages of acquisition

In Chapter 3, we pointed out how little of this gradual growth of language ability can be attributed to teaching. To be sure, adults and even older children will teach individual words. (But one an hour? I doubt it.) In addition, adults tend to speak to children more clearly and in simpler sentences than they use with other adults. So to some extent, children don't have to deal with the full daunting complexity of the language all at once.

We also noted, though, that children get very little grammatical correction, and are liable to ignore or resist correction when it does take place. Here's another famous example, cited by Martin Braine.

CHILD: Want other one spoon, Daddy.
FATHER: You mean, you want the other spoon.
CHILD: Yes, I want other one spoon, please Daddy.
FATHER: Can you say "the other spoon"?
CHILD: Other . . . one . . . spoon.
FATHER: Say "other."
CHILD: Other.
FATHER: "Spoon."
CHILD: Spoon.
FATHER: "Other spoon."
CHILD: Other . . . spoon. Now give me other one spoon?

Knowing more than they say

Impressionistically, one-and-a-half-year-old children understand an amazing amount of what you say to them—even if their speech consists only of one-word utterances and their spoken vocabulary contains only fifty words. That is, their comprehension is way ahead of their production. Some rather simple experiments show this in striking fashion.

First, consider their phonology. Where I have written whole words in transcribing babies' utterances above, I was not showing you any of the simplification that they wreak on the pronunciation. Typically clusters of consonants may be simplified ("spoon" is pronounced "poon") and final consonants may be omitted ("bus" is pronounced "buh").

There are other cases that are even more intriguing. The linguist Neil Smith writes that his son consistently substituted "f" for "th," so that "thick" came out "fick." But it wasn't that he couldn't pronounce "th," since at the same time he used "th" instead of "s," so that "sick" came out "thick"! So there is evidently some system in place that goes beyond motor control alone.

A final complication is that children often don't hear what they're doing. If you deliberately pronounce a word the way your child does, he or she will get mad at you and tell you to say it right. If you tell your child to say "duck", not "guck," most of the time you'll get "guck" and a blank stare. Perhaps it's like not being able to hear your own accent.

Knowing more than they say

Such phonological facts are easily observable. Testing syntactic understanding takes more sophisticated tests. It can be shown, though, that children as young as the one-word stage (seventeen months, say) appreciate some of the subtleties of syntactic structure.

Here's one kind of experiment, developed by Kathy Hirsh-Pasek and Roberta Golinkoff. Let's sit a very young child down in front of two side-by-side TV screens. The left-hand screen shows, say, Big Bird tickling Cookie Monster; the right-hand screen shows Cookie Monster tickling Big Bird. And out of a loudspeaker between the two screens, a voice says, "Look! Big Bird is tickling Cookie Monster!" (We have already made sure the child can identify Big Bird and Cookie Monster.) What happens? It turns out that the child will look much longer at the left-hand screen, which correctly depicts what the sentence describes. That is, the child appreciates the fact that in English the actor reliably precedes the verb and the patient follows it. Remember, not all languages have this order, so the child has to have learned something about *English*. And this effect can be observed as young as seventeen months—in many children barely the onset of the production of one-word utterances.

Knowing more than they say

Suppose I hand you a doll and say one of the sentences in (1). (I'll use all capitals so as not to bias your interpretation.)

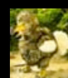
(1) *a* This is DAX.
b This is a DAX.

In the first case you will probably take "DAX" to be the name of the doll; in the second, because of the indefinite article "a," you will probably take "DAX" to be a word for doll or for some special kind of doll.

Children in the one-word stage—again as early as seventeen months—can be shown to know this too. How? We take the doll back and put it with a bunch of other things—blocks, toy cars, and, crucially, another doll. Then we say:

(2) *a* Could you give me DAX?
or
b Could you give me a DAX?
(depending on whether we first said (1a) or (1b))

In the a. case they will tend to hand you the same doll you gave them in the first place—that is, the doll named DAX. But in the b. case they are as likely to give you one doll as the other: you are asking them for any old DAX. So they evidently know that an indefinite article signals a common noun, and its absence signals a name—more than a year before they will be using indefinite articles themselves.



This is an important key to how they can learn all those words without being taught: using a combination of their understanding of the context in which a sentence is uttered plus the syntactic pattern of the sentence, they can often formulate fairly precise guesses about the meanings of unknown words.

Getting to children's knowledge

How can we tell whether children are using a rule of mental grammar in *producing* speech—how do we know they're not just imitating what they've heard? One way is by observing them saying things they've never heard.

EXPERIMENTER: This is a wug. [Pointing to a cartoon of a cute little bird-like object.] Now there is another one. There are two of them. There are two . . .
CHILD: Wug.

This child obviously doesn't know the pattern for forming plurals yet. Very roughly, about three-quarters of the four- and five-year-olds tested gave the correct answer "wugs" (with the ending pronounced z), while nearly all the six- and seven-year-olds got it right. Since they couldn't ever have heard the words "wug" and "wugs" before, they had to use a principle of mental grammar to construct the latter from the former. And we see that this rule is not reliably available for speech production till the age of six or so.

Getting to children's knowledge

Discussion question: relation with numerical cognition (Emily)

Discrete vs. continuous substances = failure from syntactic cues

"Can you hand me the wugs?" [discrete]

"Can you hand me the wug?" [continuous]

Since children are sensitive to these cues, what does it mean that they fail? Might there be other causes? How might these relate to numerical cognition?

Getting to children's knowledge

Another way to discover a child's mental grammar is to observe *systematic mistakes*: things the child says that show a consistent pattern different from adult speech. For instance, in learning to form wh-questions, children often go through a number of different stages. Here are some samples, reported by Edward Klima and Ursula Bellugi.

Stage 1 (around two and a half years):
What book name?
Why you smiling?
What soldier marching?
Stage 2 (around three and a half years):
What he can ride in?
Which way they should go?
Why kitty can't stand up?
Stage 3 (around five):
Where will you go?
Why can't kitty see?
Why don't you know?

Stage 1
No the sun shining.
No a boy bed.
No sit there.
Stage 2
He no bite you.
I no want envelope.
I no taste them.
Stage 3
I didn't did it.
You didn't caught me.

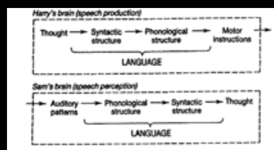
Stage 1 walked, played, came, went
Stage 2 walked, played, comed, goed, holded
Stage 3 walked, played, camed, wented
Stage 4 walked, played, came, went, held

Main points

1. Children understand a great deal more than they can imitate, showing that they have constructed grammatical patterns.
2. Children *don't* just imitate what they've heard. They are always coming up with novel utterances, which are patterned—implying that they have a mental grammar.
3. The patterns of their utterances are to some extent stripped down from the adult patterns, in particular leaving out function words and inflections, and shortening utterances to within narrow limits of a few words.
4. BUT—Their patterns have their own life, a life that cannot be induced from the input.

Grammar!

Organization of mental grammar



Functionalism

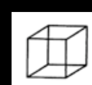
To help us see what functionalism is about, let's think about the videotape again. In order to store TV pictures, a videotape must carry a code that expresses certain distinctions, and this code must be stored in terms of basically one-dimensional patterns of magnetization on a tape. So we can ask how the code could be organized so that the videotape can do its job. As a pattern, it doesn't matter too much whether we put the code on a magnetic tape or on something else comparably one-dimensional, say a punched paper tape or a bar-code that can be read by an optical scanner: it's the pattern that counts. Similarly, we can study the patterning of speech sounds—their order, the differences and similarities among them, and their contributions to understanding—to a certain degree independently of the neural medium in which they are physically encoded.

Finding out properties of the grammar

What is a linguistic experiment? As in other sciences, the strategy is to study unobservable phenomena by relating them to things that *are* observable. If we want to measure the mass of an electron or the sun, we can't just weigh them on a scale. We have to use some sort of indirect means to get at what we want to measure—we have to think of something else we *can* measure that is connected to what we really want to know, in what we think is a reliable way. The same is true with the mental principles behind language. The only difference is that linguistic experiments have to do with the inside of our heads instead of external objects.

That's all there is to it. The idea is that although we can't observe the mental grammar of English itself, we *can* observe the judgments of grammaticality and meaning that are produced by using it.

Visual & linguistic experiments



a Visiting relatives can be boring.
b Relatives who are visiting you can be boring.
c Going to visit relatives can be boring.

Figure 4.9 *Three impossible figures*

Parallel to the impossible visual examples in Figure 4.9 are all the ungrammatical sentences of the last chapter, such as "An oboe not is octopus an." We know immediately, intuitively, that there is something the matter with them. ("Intuitive" judgments are just judgments that follow from unconscious principles. We make the judgment but can't say exactly why.)

These experiments are so simple and reliable that all we have to do is present them to observers and ask them what they see. Moreover, it is clear that our judgments of these figures have nothing to do with what we were "taught about seeing"; that these judgments require no conscious thought; and that at the same time it's very hard to be explicit about why the figures look the way they do. That is, these visual judgments have all the same symptoms as judgments about sentences. I'm suggesting, then, that the two kinds of judgments have similar status as experimental evidence.

Back to the functionalist approach

Drawing this all together: the functionalist approach to mental grammar is to make experimentally testable hypotheses about the organization of information and knowledge in the brain, without too much concern for the moment about how the brain physically encodes this information.

[redacted] an understanding of the functional organization of language would inform research into how the brain encodes information: however the brain works, it must be able to encode information with *these* sorts of patterns.

In a functionalist theory, what does it mean to say we have a certain principle in our mental grammars, as part of the equipment we bring to understanding and creating novel sentences of English? Let's take a very simple principle of English—for inltance, that the subject of a sentence (normally) precedes the verb. If this principle is somehow in our heads, then the terms "sentence," "subject," and "verb" must be too. What does it mean to say we unconsciously know and use these terms?

Suppose we think of our stored knowledge as the contents of some curious sort of filing cabinet in the brain. The information in the filing cabinet isn't stored in a form readable by us outside observers. Why should it be? It's not there for the benefit of outside observers, it's there for the use of the rest of the brain.

[redacted] We use a term like "sentence" to distinguish a particular class of word *sequences* from everything else; the claim is that the brain makes the same distinction. We use terms like "subject" and "object" to pick out word sequences that are particular *parts of* sentences; the claim is that the brain can pick out the same parts. Finally, the whole condition, "the subject of a sentence precedes the verb," states a relation among various parts of the sentence; the claim is that the brain—however it identifies and stores these parts—imposes the same relation.

Speech perception, speech production, & phonology

One of the primary intuitions we have about language is that it comes divided into words, and that the words can be neatly divided into syllables and individual speech sounds. The phonological structure of language is an encoding of this sequence of sounds. It turns out that this sequence is a considerable abstraction of what physically takes place in speech. The acoustic stream we hear as speech shows no such neat divisions. To understand why, it's useful to see how speech is produced.

For a convenient analogy, think about how a trumpet works. The player presses his or her lips together and forces a stream of air through them, producing a sort of buzzing sound. When the trumpet is placed against the vibrating lips, the air column in the trumpet is forced into vibration as well. The way the air column vibrates is a function of the vibration of the lips interacting with the resonant frequencies of the tube; the tone quality we hear has to do with which natural harmonics of the tube are produced in what proportions. If we change the tighness of the lips, the pitch of the whole system changes; if we place a mute in the trumpet's bell, we change the mix of resonant frequencies of the tube, and so the tone quality changes.

Now let's imagine a trumpet whose tube is made out of rubber instead of brass, so it can be stretched and pinched in various ways.

Speech perception, speech production, & phonology

Let's also imagine that a second tube branches off somewhere in the middle, and that air can be directed out of either tube or both. Such a horn probably won't have the clear resonances of a regular trumpet, but it will be able to produce a much greater variety of tone colors, because its resonant frequencies can be altered so much.

Next let's shove this trumpet down the player's throat.

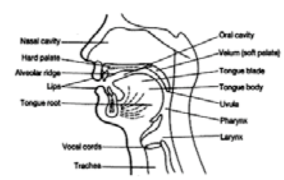


Figure 5.1 *The vocal tract*

Speech perception, speech production, & phonology

During speech, the movements of the vocal tract are smooth and continuous. For instance, try saying the word *wow*, and pay attention to what your mouth is doing. You don't hold a *w* (or *oo*) sound, with pursed lips, then suddenly switch to an *ah* sound, with mouth and lips open, then instantaneously switch back to the *w*. Trying say it that way: oo-*ah*-oo, with sharp transitions between the sounds. You can hear how unnatural it is.

As a result of the smooth transition between positions of the vocal tract, the acoustic signal produced by the vocal tract also shows a smooth transition from one sound to the next, without any abrupt boundaries. Consequently, the signal that the hearer *perceives* as neatly divided speech sounds is actually far from it. The waveform undergoes continuous change as a result of continuous change in the shape of the vocal tract.

Nevertheless, in order to structure speech, it is necessary for the brain to code the sequence of speech sounds and their combinations into words and sentences. This code of speech sounds is referred to as *phonological structure*. In this section we have seen that phonological structure is a distinct kind of mental organization from either vocal tract instructions or auditory patterns (see Figure 5.2).

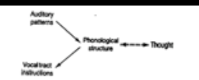


Figure 5.2 *The place of phonological structure in the information flow of language*

Speech perception, speech production, & phonology

To make this more vivid, think about the divisions between words. In writing, of course, we leave spaces between words, and this mirrors our perception: we (almost) never have trouble hearing where one word ends and the next begins. But listen to yourself say the following pairs of sentences:

(1) *a* I don't really think it's a parent.
b I don't really think it's apparent.
(2) *a* Have you looked at this guy yet?
b Have you looked at the sky yet?
(3) *a* We needed a cantor.
b We need a decanter.*

In case these examples from English don't convince you, think about the last time you heard people speaking a language you didn't know. Could you tell where one word ended and the next began?

A profound point is lurking here. We have discovered that the words we consciously hear and pronounce are not in any physical sense "out in the world." We can't find them through physical measurement of acoustic waveforms. Rather, the way we experience the stream of language seems to have more to do with the patterns of phonological structure in our heads, in which the speech sounds and words are clearly demarcated.

Speech perception, speech production, & phonology

Intuitively, speech sounds are unitary: the sound *t* is just that, a single sound. However, one of the major discoveries of phonological theory, originally developed by Nicholas Trubetzkoy and Roman Jakobson in the 1920s and 1930s, is that speech sounds are encoded in the brain in terms of more primitive specifications called the *distinctive features* of speech sounds. Some of these features are listed in (4). If you pay very careful attention to what is going on in your mouth and throat as you make different sounds, you can verify the features for yourself.

a Significant constriction of the vocal tract (consonants)
vs.
Vocal tract unconstricted (vowels)
b Vocal cords tensed and therefore in vibration (sounds such as *b, g, d, z, n, th* as in "there," all vowels)
vs.
Vocal cords relaxed and therefore not vibrating (*p, t, k, ch, f, sh, th* as in "thin")
c Velum lowered so air passes through nose (*m, n, ng* as in "sing," vowels in French "on" and "fin")
vs.
Velum raised so air passes only through mouth (all other sounds)
d Air flow through mouth completely blocked (*p, b, m, t, d, n, k, g, ng, ch, j*)
vs.
Air flow through mouth not completely blocked (all other sounds)
e Most constricted part of mouth at lips (*p, b, m, f, v, w, u*)
vs. at tip of tongue (*t, d, n, s, z, l, th*)
vs. at body of tongue (*ch, sh, j*)
vs. at back of tongue (*k, g, ng*, German *ch* as in "Bach")

Speech perception, speech production, & phonology

The relations among sounds provided by distinctive features enable us to explain many curious aspects of pronunciation. For a simple example, the plural suffix for English nouns is pronounced three different ways: as a *z* sound in words like "dogs," as an *s* sound in words like "cats," and as an *uhz* sound in words like "horses." (It is always *spelled* "s" or "es"; but I want to pay attention to how it is actually *pronounced*. For simplicity, I am disregarding irregular plurals like "oxen," "mice," and "sheep.") What decides among these three choices in any given word?

a If the noun ends with one of the sounds *s, z, sh, ch, j*, the plural is pronounced *uhz*.
b If the noun ends with one of the sounds *p, t, k, f*, or *th* as in "death," the plural is pronounced *s*.
c If the noun ends with anything else, the plural is pronounced *z*.

These classes of sounds may look arbitrary, but in fact distinctive-feature analysis shows us the method behind the madness. The class in (5b) consists of sounds that are "unvoiced"—that is, in which the vocal cords are not vibrating. The class in (5c) consists entirely of sounds that are "voiced"—in which the vocal cords vibrate. Two of the pronunciations of the plural, *s* and *z*, are articulated identically except for this feature of voicing; we can see the pronunciation of the plural as "agreeing" in voicing with the end of the word it is attached to.

What about the class in (5a), which includes both voiced sounds (*z* and *j*) and unvoiced ones (*s, sh*, and *ch*)? This class includes sounds whose articulation is very close to that of the plural ending: *s* and *z* themselves plus the sounds articulated with the body of the tongue. Here a vowel sound is inserted before the plural ending to prevent interference between the two consonants—and the plural itself is pronounced *z* to agree with the voicing of this vowel.

Speech perception, speech production, & phonology

Not only can the choice among the three pronunciations of the plural be predicted from the features of the final sound of the word, but the choice is *productive*—that is, we can produce plurals for new nouns on the spot. We don't just memorize the plural ending for every noun we know (though we probably *do* memorize the irregular cases). To see this, suppose I introduce new words to you and ask you what their plurals are. If you had to learn the plural form by memorization, you couldn't carry out this task. Here are some words (they happen to be borrowed from Yiddish).

a kvetch …es
b dybbuk …s
c shmeggeggie …z

There is a great deal more to phonological structure than just distinctive features. In addition to a sequence of speech sounds, words and phrases carry stress, or relative emphasis among their syllables. In many languages (such as Chinese, Vietnamese, and many West African languages), words carry with them an inherent melody (or sequence of tones) that is as much a part of the word as its speech sounds. Words and phrases also have an inherent rhythm that determines their temporal flow. For each of these aspects of the sound structure of language there is a growing literature of analysis and theory, with much lively dispute.

Speech perception, speech production, & phonology

Meanwhile, on to further complications. Let's return to speech perception—how the brain gets from the activation of auditory neurons to a perception of discrete speech sounds. We've already noted how smeared the auditory input is, and how this presents a severe problem for perception, which has to locate the boundaries between speech sounds and even between words. But that isn't the only problem that speech perception faces.

the process of auditory perception analyzes the acoustic signal into three separate but simultaneous factors: *who* is speaking (voice recognition), *what* the speaker is saying (language perception), and *how* it is being said (the speaker's tone of voice or emotional affect).

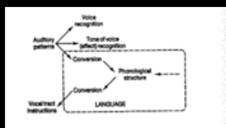


Figure 5.3 The auditory signal feeds language and two other specialized processors

Syntactic structure

a An X is not a Y.
b Since an X is not a Y, a Z is not a W.
c X Verbs that S.

What fits into the slots marked "X," "Y," "Verb," and "S" in these patterns? We can't describe these just in terms of their sound. Rather, we need the notion, familiar from traditional grammar, of "parts of speech" such as noun, verb, adjective, and preposition—plus ways of combining them. *X, Y, Z,* and *W* in patterns (1a) and (1b) have to be filled by nouns such as "nunnery," "banana," "oboe," and so forth; *Verb* in pattern (1c) has to be filled by a verb such as "thinks," "believes," "expects," and so forth.

There is a further complication in pattern (1c). As we saw in Chapter 2, *S* has to be filled with another pattern, a sentence that can stand on its own. In the examples in Chapter 2, *X* was filled by a name (or proper noun) such as "Larry," "Moe," or "Curly." But it can also be filled by a larger pattern of words consisting of a common noun and a collection of modifiers—a so-called *noun phrase*. In the sentences in (2), I've underlined the noun phrase that takes the place of *X* in pattern (1c), and I've marked in bold the *head noun*, the noun that everything else modifies.

a The big black bear thinks that you won't shoot him.
b A woman in the lobby with a book under her arm believes that an oboe is not an octopus.
c The tall boy who Bill met yesterday expects that the world economy will disintegrate within a year.

It should be fairly obvious that these syntactic categories and patterns—noun, verb, noun phrase, and so forth—can't be characterized in terms of phonological structure. For one thing, the very same sequence of sounds can serve as different parts of speech, as seen in the sentences in (3); you can doubtless multiply examples *ad libitum*.

a We're going to rock around the clock. (rock = verb)
b We put some rock around the clock. (rock = noun)
c Beth threw the ball. (threw = verb)
d It went through the window. (through = preposition)

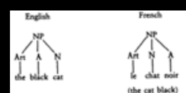
Syntactic structure

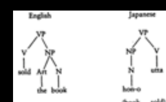
Syntactic structure is closer to meaning than sound is—it's the last way-station enroute from sound to meaning—so it strongly reflects certain aspects of meaning. But, as I want to show, there are other properties of syntactic structure that don't have much to do with meaning. Rather, they have to do with organizing the elements of meaning into linear order so that they can be pronounced, and at the same time marking the relations among these elements so that they can be re-identified by the hearer.

If parts of speech don't have to do with meaning, what *do* they have to do with? It should be evident by now that the classification of words into parts of speech determines their roles in patterns.

a noun can appear with a plural ending: "dogs," "bananas," "earthquakes." A verb, on the other hand, can appear with a past tense ending: "helped," "believed," "procrastinated." Notice that our ability to use these endings is *syntactic* knowledge, and doesn't follow from the meanings of the words. In terms of meaning, it would make sense if nouns that name actions could appear with a past tense. But there are no words "earthquaked" or "concerted" which name an earthquake or a concert that occurred in the past. Likewise, in terms of meaning, it would make sense to be able to put a plural ending on a verb to mean that the action was performed more than once. But we can't say "Bill will dances" to mean he will dance several times. (The "-s" ending in "Bill dances", of course, indicates not plural, but that there is a third-person-singular subject.) In other words, the availability of past tense and plural endings correlates with the syntactic distinction between verbs and nouns, not with the distinction in meaning between objects and actions.

Syntactic structure





Ambiguity from structure to meaning

a Many people here have read two books.
b There are two particular books (say, *Gone with the Wind* and *Fear of Flying*) that many people here have read.
c Many people here have read two books, but not necessarily the same two.

Levels of Representation
Marr (1982)

Describing vs. Explaining

As one reflected on these sorts of issues in the early 1970s, it gradually became clear that something important was missing that was not present in either of the disciplines of neurophysiology or psychophysics. The key observation is that neurophysiology and psychophysics have as their business to *describe* the behavior of cells or of subjects but not to *explain* such behavior. What are the visual areas of the cerebral cortex actually doing? What are the problems in doing it that need explaining, and at what level of description should such explanations be sought?

And present is a clear understanding of what is to be computed, how it is to be done, the physical assumptions on which the method is based, and some kind of analysis of algorithms that are capable of carrying it out.

This was what was missing—the analysis of the problem as an information-processing task. Such analysis does not usurp an understanding at the other levels—of neurons or of computer programs—but it is a necessary complement to them, since without it there can be no real understanding of the function of all those neurons.

But the important point is that if the notion of different types of understanding is taken very seriously, it allows the study of the information-processing basis of perception to be made *rigorous*. It becomes possible, by separating explanations into different levels, to make explicit statements about what is being computed and why and to construct theories stating that what is being computed is optimal in some sense or is guaranteed to function correctly.

The three levels

Computational theory	Representation and algorithm	Hardware implementation
What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?	How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?	How can the representation and algorithm be realized physically?

Although algorithms and mechanisms are empirically more accessible, it is the top level, the level of computational theory, which is critically important from an information-processing point of view. The reason for this is that the nature of the computations that underlie perception depends more upon the computational problems that have to be solved than upon the particular hardware in which their solutions are implemented. To phrase the matter another way, an algorithm is likely to be understood more readily by understanding the nature of the problem being solved than by examining the mechanism (and the hardware) in which it is embodied.

A computational example

For example, the Arabic, Roman, and binary numeral systems are all formal systems for representing numbers. The Arabic representation consists of a string of symbols drawn from the set (0, 1, 2, 3, 4, 5, 6, 7, 8, 9), and the rule for constructing the description of a particular integer *n* is that one decomposes *n* into a sum of multiples of powers of 10 and unites these multiples into a string with the largest powers on the left and the smallest on the right. Thus, thirty-seven equals $3 \times 10^1 + 7 \times 10^0$, which becomes 37, the Arabic numeral system's description of the number. What this description makes explicit is the number's decomposition into powers of 10. The binary numeral system's description of the number thirty-seven is 100101, and this description makes explicit the number's decomposition into powers of 2. In the Roman numeral system, thirty-seven is represented as XXXVII.

If one chooses the Arabic numeral representation, it is easy to discover whether a number is a power of 10 but difficult to discover whether it is a power of 2. If one chooses the binary representation, the situation is reversed. Thus, there is a trade-off; any particular representation makes certain information explicit at the expense of information that is pushed into the background and may be quite hard to recover.

This issue is important, because how information is represented can greatly affect how easy it is to do different things with it. This is evident even from our numbers example: It is easy to add, to subtract, and even to multiply if the Arabic or binary representations are used, but it is not at all easy to do these things—especially multiplication—with Roman numerals. This is a key reason why the Roman culture failed to develop mathematics in the way the earlier Arabic cultures had.

Cash register computation

The most abstract is the level of *what* the device does and *why*. What it does is arithmetic, so our first task is to master the theory of addition.

Addition is a mapping, usually denoted by +, from pairs of numbers into single numbers; for example, + maps the pair (3, 4) to 7, and I shall write this in the form $(3 + 4) \rightarrow 7$. Addition has a number of abstract properties, however. It is commutative: both $(3 + 4)$ and $(4 + 3)$ are equal to 7; and associative: the sum of $3 + (4 + 5)$ is the same as the sum of $(3 + 4) + 5$. Then there is the unique distinguished element, zero, the adding of which has no effect: $(4 + 0) \rightarrow 4$. Also, for every number there is a unique "inverse," written (-4) in the case of 4, which when added to the number gives zero: $[4 + (-4)] \rightarrow 0$.

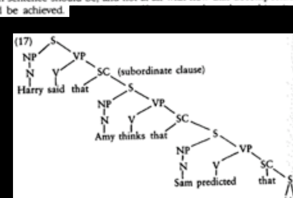
Notice that these properties are part of the fundamental *theory* of addition. They are true no matter how the numbers are written—whether in binary, Arabic, or Roman representation—and no matter how the addition is executed. Thus part of this first level is something that might be characterized as *what* is being computed.

This whole argument is what I call the *computational theory* of the cash register. Its important features are (1) that it contains separate arguments about what is computed and why and (2) that the resulting operation is defined uniquely by the constraints it has to satisfy.

Mapping the framework: Computational Theory of Language

Chomsky's (1965) theory of transformational grammar is a true computational theory in the sense defined earlier. It is concerned solely with specifying what the syntactic decomposition of an English sentence should be, and not at all with how that decomposition should be achieved.

(17)



A question

Discussion question: levels of hierarchy and Jackendoff (Erin)

Is it possible to work on the levels simultaneously? Jackendoff seems to propose this for language. Are there any pitfalls that would be associated with this?

Mapping the Framework: Algorithmic Theory of Language Learning

The "how" of language learning: want computational-level description of the problem (word segmentation, speech perception, word learning, sentence structure, metrical stress, etc) and what the algorithm is that a learner could use to solve it (input, output, and process that takes input to generate output)

Considerations: input available, hypotheses available for generating output, psychological plausibility of algorithm

Framework for language learning

What are the hypotheses available (for generating the output)?

Ex:

Adjective before noun (ex: English)
Red apple

Noun before adjective (ex: Spanish)
manzana roja (apple red)

What data is available, and should the learner use all of it?

Ex:

Ignore special use of adjective before noun in Spanish
If the adjective is naturally associated with the noun:
la blanca nieve (the white snow - snow is naturally white)

How will the learner update beliefs in the competing hypotheses?

Ex:

Probabilistic update, based on data intake (Bayesian, Linear reward-penalty)