CAS LX 522
Syntax I

1

Introduction to the enterprise

Once upon a time...

Snoopy kissed whats-her-name after Pigpen chased an orange thing.

Who's the girl that Snoopy kissed after Pigpen chased the orange thing?

What's the thing that Snoopy kissed Lucy after Pigpen chased?

What's the thing that Snoopy told Lucy that Pigpen chased?

English vs. word salad

1) Chris looked over the numbers.
2) Pat peeked over the fence.
3) Chris looked the numbers over.
4) Pat peeked the fence over.

Distiguishing English from not-so-English

1) Slept cat the.
2) Slept the cat.
3) The slept cat.
4) The cat slept.
5) Cat the slept.
6) Cat slept the.

When presented with a sequence of English words, any (native English) speaker can tell you whether it makes a sentence of English. — *How?*

Knowledge of language

Native speakers "just know" what is part of their language.

But it's *tacit* knowledge—we can't just explain what it is that makes a sentence English. It just is. Or isn't.

Our task: Exploring and characterizing what the nature of this knowledge is, and how it differs between languages.

What speakers know

Although we can't explain our own knowledge of our native language, we can deduce it—and, in simple cases, we have a kind of intuition about what it is.

1) The cat slept.
2) Slept cat the.

The noun

Cat the does not make a good subject of a sentence, it has to be *the cat*.

In fact, *the* can't really stand anywhere except before a noun.

So, we hypothesize that English speakers know something like a general rule: *the* comes before nouns.

Subject verb

The noun can be the subject of a sentence.

And the subject seems to come before the verb.

So, we hypothesize that English speakers know something like a general rule that subjects of a sentence come before the verb.

Formalities

We can make these hypotheses formal, explicit:

A subject is made of *the* and a noun.

A sentence is made of a subject and a verb.

S → subject V

subject → the N

N → cat

V → slept

Have we done it?

Perhaps that's it, perhaps we have now described English. Let's see.

There are lots of other nouns. *Dog* for example. And there are other verbs too. Like *coughed*.

If these rules describe English, then *The dog coughed*, *the cat coughed*, and *the dog slept* should be judged to be English sentences.

Hooray!

This is exciting! Maybe we *have* done it!! If this is what English speakers know about English, then all and only the sentences *generated* by these rules should be judged as English.

Hooray! Er...

This is exciting! Maybe we *have* done it!! If this is what English speakers know about English, then all and only the sentences *generated* by these rules should be judged as English.

1) The dog chased the cat.

Hmm.

What went wrong?

Although *the dog chased the cat* is judged to be English, our rules do not generate this sentence.

Just looking at it, we can see that the problem is that some verbs, like *slept* and *coughed* describe something performed by just one individual, but *chase* is something one individual does to another.

Subject Verb Object

New rule: A verb with both a do-er ("subject") and a do-ee ("object") (let's call such verbs "transitive") comes between them.

S → subject V

subject → the N

S → subject Vt object

object → the N

N → cat, dog

V → slept, coughed

Vt → chased

Subjects and objects

We notice that our "subject" and "object" rule both look the same.

Also, notice that we can also say *A dog chased a cat*. So, *a* and *the* are probably the same kind of thing. We'll call them "determiners" (though you might have called them "articles").

Probably anything that can be a subject can also be an object. So we can simplify our rules.

"Nouny" phrases

We need a name for these "*the noun*" things. More than one word (a *phrase*, if you will), where the noun seems like the most important part. How about "noun phrases"?

S → NP V

NP → Det N

S → NP Vt NP

N → cat, dog

Det → a, the

V → slept, coughed

Vt → chased

&c

Right.

- The grumpy cat chased the terrified dog.

Ah. So, NPs can have *adjectives* like "grumpy" and "terrified" in them.

Ok, so our "theory of English knowledge" is still insufficient, but there's a fairly clear way to extend it.

Grammar

Supposing that we finally get to the end of this procedure, what we will have constructed is a *grammar*—a system that can distinguish strings of words into "English" and "not English."

The sort of grammar we've been constructing is a *generative grammar*. The theoretical claim is that all —and only—strings that it generates will be judged by native English speakers as being "English."

It is a theory—or a model—of what English speakers know about English.

The S→NP V neuron?

This is *not* a claim that the *actual* rules we're coming up with are somehow encoded in the brains of native English speakers.

The system we're hypothesizing *characterizes the knowledge*, but who knows how the neurons are organized.

We can still learn a lot about the structure of language though—and maybe learn what kinds of things to look for among the neurons.

Intuitions

The primary thing we're trying to explain is why people have the *intuitions* they have about language.

For a given string of words, a native speaker can say whether it is part of their language. But probably can't tell you *why*.

These intuitions are quite stable across speakers. We seek the basis for these intuitions.

Stars

The notational convention for marking a sentence that is not part of the language is putting an asterisk ("star") in front of it.

1) *Cat the slept.
2) The cat slept.

Unacceptability

A string (of words) can be unacceptable for a number of different reasons. Some are important for building our model, and some are not.

1) *Big that under staple run the jump swim.
2) My toothbrush is pregnant again.
3) The rat the cat the dog chased caught escaped adeptly.

Ambiguity

Frequently, a sentence can be used to express more than one meaning.

1) I walked on the bank.
2) You can't stop a philosopher with a thesis.
3) All doors will not open.
4) Nothing cleans better than Bio-Soy Orange™.

*Ambiguity

But sometimes an ambiguity we might have expected to find is not there. So the *sentence* isn't ungrammatical, but one intended *interpretation* is.

1) How did John say that Mary used the jackhammer?
   Answers: a) Incompetently, b) Quietly.
2) How did John ask if Mary used the jackhammer?
   Answers: a) *Incompetently, b) Quietly.

But I would never say that

It generally does not matter that you'd never (or almost never) *use* a sentence being judged.

Given the circumstances under which the sentence would be appropriate (rare as they may be), would it be English?

Quite often one needs to construct rather artificial sentences to test specific hypotheses.

Felicity

Another kind of unacceptability, which we generally won't be concerned with in this class, is the kind that arises from a mismatch with the preceding discourse.

1) Who bought the lamp?
2) #Noel bought *the lamp*.

We care about this when studying constraints on the structure of discourse, but here we're studying constraints on the form of sentences in isolation.

Mysteries

Our knowledge of language is very complex, but not available to introspection.

Children acquire language very quickly, in the same way, and with a stable result.

Different languages turn out to have a lot in common—there are a lot of possible properties language might have, but yet never seems to.

How did we get this?

Children certainly are not told things like "Billy, if I *ever* catch you using a subject pronoun that matches the reference of a proper name object, you'll be eating only asparagus for a week!"

Yet they know that *He saw John in the mirror* can't mean that John saw himself.

And—really—what *could* they hear that would teach them this?

Kids don't just imitate

Imitation by itself could never work as a means of acquiring this kind of language knowledge.

The knowledge we're talking about differentiates English sentences from non-English sentences. (Or whatever language)

Which ones are good? The ones you hear?

- The first platypus to eat twenty-five maroon gummi bears will win a prize.

Things adults don't say

Plus, kids say things they've never heard an adult say.

- Me playing.
- What do you think what the puppet has eaten?

How would this come about?

Sure, ok. You generalize.

So, maybe kids hear what parents say and recognize the patterns, and come up with general rules.

That's not so hard, is it?

Let's try it out.

Making questions

1) Sue should borrow my guitar.
2) Sue borrowed my guitar.
3) Should Sue borrow my guitar?
4) Did Sue borrow my guitar?
5) What should Sue borrow?
6) What did Sue borrow?

That's easy

Yes-no questions are formed by taking the second word and putting it in the front. Unless there's no word like *should*, then you just put *did* in the front, and use a bare (untensed) verb.

Wh-questions are formed by removing something and putting *who* or *what* in front of the yes-no question form.

See? Simple.

- My roommate should borrow my guitar.
- Roommate my should borrow my guitar?
- Roommate my did borrow my guitar?
- Who did borrow my guitar?
- What roommate my should borrow?
- Why person any would think this is hard?

Trying something else

- Sue said that Mary borrowed my guitar.
- Sue said Mary borrowed my guitar.
- What did Sue say that Mary borrowed?
- What did Sue say Mary borrowed?
- Who did Sue say borrowed my guitar?
- Who did Sue say that borrowed my guitar?

And one more

- Mary saw her in the mirror.
- Mary saw her duck in the mirror.

Why can't *her* be *Mary*? (Except if it's Mary's duck—but it can't be Mary who is ducking). How is a kid supposed to deduce this?

Poverty of the stimulus

The point is: The linguistic input that a child gets is insufficient to determine which of the possible rules of grammar are the right ones.

Yet, children always acquire the same rules.

1, 2, 3, —, —, —?

4, 5, 6? — 5, 7, 11? — 5, 8, 13? — 3, 2, 1?

What are we left with?

Of course, you know where this is going: Having language = being human.

Rocks, ferns, cats, apes don't soak up language when surrounded by it.

Birds have wings, people have arms.

What determines whether you're a bird is whether your parents are birds.

Human language

The point: What makes an organism a human is something about what is encoded in the genes.

Only humans have language.

So *something* that makes language possible must be encoded in the genes.

Universal Grammar

This is the idea of Universal Grammar (UG), which we take to refer to the human-specific cognitive structures underlying language.

Languages differ, too. So, apart from the (species-)universal *principles* of language, there are differences in how they interact and operate. *Parameters* of variation.

Parameters

Basic word order:

English (SVO): Akira bought a book.

Japanese (SOV): John ga hon o katta.

Question formation (where *what* goes):

What did Akira buy?

John ga nani o katta no?

Explanation

If languages are all governed by the same principles (which children don't have to learn), then the child's task is really to:

Learn the pronunciations and meanings of the words.

Determine the "settings" of the parameters for the language they are acquiring.

What are the principles, parameters?

This is the "big picture" take on what we're trying to accomplish in Syntax.

How can languages vary?

In what ways don't languages vary?

Do certain language properties "group" together? Could they depend in some way on the same parameter?

Prescriptive rules

We no doubt remember being taught things like this:

1) A preposition is something you should never end a sentence with.
2) It is important to always avoid splitting an infinitive.

But—there's a reason why these were painful to learn. They *aren't* rules of English.

Prescriptive rules

Prescriptive rules are generally somewhat arbitrary. Somebody's idea of what the language "ought" to be like, or hanging on to how the language used to be even after it has changed.

If these were actually rules of English, they wouldn't need to be taught (to native speakers at least).

Mainly, they serve as a "secret code" that educated people use to identify each other.

Where is English?

When we speak of "English," what are we referring to?

Every native speaker has a complete knowledge system of their language.

As far as the grammar is concerned, it's all part of a native speaker's cognitive makeup. (Vocabulary is a different thing...)

I-language, E-language

The notion of "English" is really an *external* notion. It's kind of an "average" of the properties of the (nearly identical) knowledge systems that the individual speakers in the community have.

What we're interested in here, in a sense, are the properties of a single speaker's knowledge of language. We might call it "English" if that speaker is part of the "English" speech community. But it's really an individual's knowledge. It's just that the community by and large has the same knowledge.

Incidentally, re: LX250

You may recall that in LX250, you did some syntax. There, you were told things like: Sentences have structures described in terms of X-bar templates (heads, complements, specifiers), there is a CP, a IP, a VP, and some NPs.

In a sense, this was kind of "skipping to the middle." We're going to back up to motivate some of these things (and argue for them), and we'll wind up with a system that is a bit different (more modern).

(So, don't just draw trees according to your LX250 rules, they won't be right. Close...but not right.)