Second language acquisition of syntax

LX 454/754: Acquisition of Syntax

Spring 2018, February 7

Scientific study of language

- What constitutes one's knowledge of language?
- How is that knowledge acquired?

Looking at adult native languages, we've found that language is very complex (cf. Syntax classes, etc.).

Looking at children, we've found that they seem to learn this complicated system with surprisingly little help from the environment.

L1 acquisition

We posited a genetic predisposition for language, something that guides the kinds of languages children learn (Universal Grammar).

- Kids learn *fast*.
- Kids end up with systems that more complicated than the input data justifies (they can judge ungrammatical sentences in the same way as other native speakers).
- Kids don't fail to learn language despite differences in environment, and without getting or making use of negative evidence.
- Kids seem to go through similar stages, across kids, across languages.

L2 acquisition

L2 acquisition seems different.

- Adults seem to have a harder time learning a second language than kids do learning their first language (there may be a "critical period").
- Adult second language learners rarely reach a native-speaker-like-level of competence.
- Adult second language learners already know a language.
- Adult second language learners are often given negative evidence ("you don't say it that way") when taught in a classroom.

Scientific study of language

- What constitutes one's knowledge of language?
- How is that knowledge acquired?

We can still study these questions in L2A as well and try to determine the answers, whether they are related to the answers we got for L1A or not.

And perhaps surprisingly, those answers might *be* related.

L2 competence

Learners of a second language have some kind of (systematic) linguistic knowledge. They have retained their L1 knowledge, and they have knowledge of a kind approximating (perhaps poorly) the knowledge held by a native speaker of the learner's L2.

This knowledge is often referred to as an *interlanguage grammar*—not L1, not L2, but something different (. . . and to what extent this knowledge might be related to or influenced by L1 or L2 is yet to be determined).

UG or not UG?

In theoretical study of L2 acquisition, a question that in some form or another has dominated the research is:

One of the big questions

How does L2A relate to UG?

What the question even means is not completely clear, but it's nevertheless gotten a great deal of attention.

UG and L2A

The basic questions are, essentially, about either the properties of the interlanguage knowledge (is it like a natural language, does it have the properties native languages have?) or the process of acquiring knowledge (is any of the "automatic" part of L1A still operative in L2A?).

However: First question depends also on how L2 interacts with L1. L1 conforms to UG, so if L2 "conforms to L1," it will also show UG-related properties.

Grammaticality judgments

One way of testing aspects of a person's L2 competence is to ask them to rate sentences in their second language.

(1) Who did you say that bought John dinner?
(1–bad 2–weird 3–natural)

(2) I wonder what will John wear tomorrow.
(1–bad 2–weird 3–natural)

Considerations on GJ tasks

As in any experiment, you may find *biases*...

- Some people are hesitant to take an extreme position, may never rate a sentence 1 or 3. (Though this turns out not to be a big factor, usually.)
- Some people may rate the sentences based on how much sense it makes, rather than on the syntactic structure. And it's hard to correct for that, because if you ask someone what's wrong with *What did you laugh after John bought for Sue?* even native speakers won't be able to say.
- It's quite possible that a given sentence might be rated badly for reasons that don't have to do with what you're testing.

Production

It's also possible to look at production data, although this is also likely to be misleading.

The biggest problem is that you cannot distinguish between "dispreferred" and "ungrammatical" in production data. Neither will appear. But each has a very different status within a theory.

This is equally true of child L1 data too, of course. It constrains the kinds of things you can really conclude from, e.g., a CHILDES search.

A case study

Kanno, Kazue (1996). The status of a nonparametrized principle in the L2 initial state. *Language Acquisition* 5(4):317–355.

Kanno identifies a particular principle of UG (the ECP—Empty Category Principle) that seems to have very different effects in the L1 and L2, and for which the effects are rather subtle (therefore untaught). She looks at whether L2'ers show evidence of knowing these ECP effects anyway.

The hope is to conclude that L2'ers *do* know the ECP, and that this knowledge must have come from UG—not the L1 (because the effects are so different), and not from instruction (because it is not taught).

The ECP in Japanese

The ECP is a principle that is hypothesized to constrain "empty" (silent) elements in the syntax. This includes traces of movement, or, relevantly here, "deleted"/"omitted" case markers.

(3) John ga sono hon o yonda.
John nom that book acc read
'John read that book.'

(4) John ga sono hon — yonda.
John nom that book Ø read
'John read that book.'

(5) * John — sono hon o yonda.
John Ø that book acc read
'John read that book.'

Complexities of the ECP in Japanese: Final particles

The basic force of the principle is that a silent element is allowed only if it can be "identified" by something else in the structure. The picky details are not crucial, but it generally separates subjects and objects—objects can be "identified" by the verb, and subjects can't.

One exception to this is when there is a final particle like *yo* in (6)—in this case, the particle can "identify" a missing subject case marker, and so the subject case marker can be left off as well.

(6) John — sono hon o yonda yo.
John that book acc read part
'John (indeed) read the book.' (I think)

Complexities of the ECP in Japanese: Topics

Another added complication is that in Japanese, either case marker can be replaced by the topic marker *wa*.

However, unlike subject markers, *wa* (even on subjects) *can* be omitted. So, this makes the pattern very subtle, since this looks a lot like an omitted subject case marker—it's different only in its interpretation.

(7) Tanaka-san (wa) itsu kaimasita ka?
Tanaka top when bought Q
'When did Tanaka buy it?' / 'As for T, when did he buy it?'

The ECP in English

The ECP is hypothesized to be a principle of UG—it constrains all native human languages. It holds of English as well, but its effects look different. English has no case markers, but it has been argued that the difference between (8) and (9) arises from the ECP.

(8) Who did you say ∅ *t* left?

(9) * Who did you say that *t* left?

The idea here is (kind of weirdly) that the "silent *that*" in (8) is able to "identify" the trace of *who* in the subject position, but the overt *that* cannot.

Kanno: How could an English speaker transfer the knowledge of these "that-trace" effects to apply them to case marker omission in Japanese?

Looking at the Japanese instruction

Having suggested that L2'ers can't get the pattern for case marker omission from the L1 (English), Kanno goes on to investigate what evidence they have gotten about case marker omission in Japanese from their instruction.

The goal is to show that they were not taught the pattern—if the L2'ers actually do respect the pattern, then it was not because they were instructed to do so. They will have "gone beyond the input"—the conclusion being that the knowledge must have originated with UG. It's part of "what languages are like."

Textbook examples of case marker omission

41 cases of object case-marker drop

(10) Enpitsu — kudasai?
pencil give
'Can you give me a pencil?'

8 cases of subject case-marker drop, in the exceptional case when it is allowed (with a final emphatic particle—these don't violate the ECP)

(11) John — sono hon o yonda yo.
John that book acc read part
'John (indeed) read the book.' (I think)

Textbook examples: dropping nominative legitimately

Certain verbs have nominative case on their objects, and case can be dropped on those objects too...

(12) John ga kankokugo (ga) dekimasu.
John nom Korean nom can-do
'John can speak Korean.'

69 of 110 such verbs in the book had the object case marker dropped.

Textbook examples: omitted subjects and initial objects

Japanese allows arguments to be omitted (somewhat like Italian *pro*-drop), so there were many cases with just one argument (the object) with no case marker.

(13) Kami — irimasu ka?
paper need Q
'Do you need paper? / Is paper necessary?'

Textbook examples: omitted topic markers

Worst of all, cases where the *topic* marker was dropped had a "very high incidence."

(14) Tanaka-san (wa) itsu kaimasita ka?
Tanaka top when bought Q
'When did Tanaka buy it?' / 'As for T, when did he buy it?'

Textbook misinformation

"*ga* [nom] might be deleted, but with a reduction of the emphasis and focus conveyed by its inclusion." (No hint that sometimes—even usually—it is not possible to leave it out.)

"If *o* [acc] is deleted, [the object] would simply lose a bit of its emphasis and focus. On the other hand, the addition of *o* would give added emphasis and focus."

Futility

There's pretty much *no way* they could have reached the right generalization based on what they were provided.

- Nom can be dropped from object position.
- Top can be dropped from subject position.
- Nom subject can be dropped with a particle.
- Explicit instruction was only about emphasis.

Experimental items

Kanno's sentences used *wh*-words. *Wh*-words in general do not allow topic marking, so if the particle is dropped from a subject *wh*-word, it could not have been a topic drop. These were the conditions.

(15) subject wa *wh*-object — verb Q?

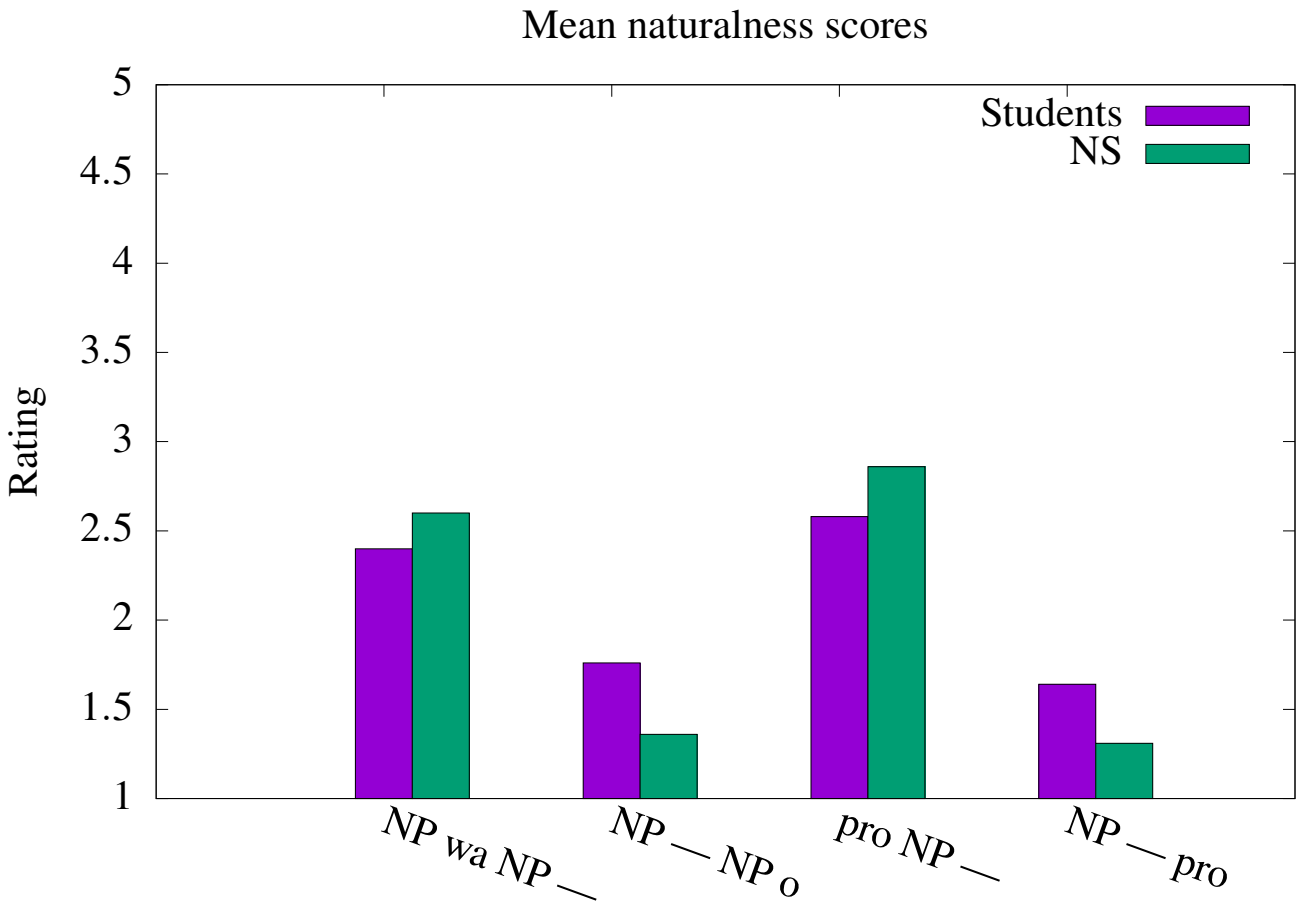
(16) * *wh*-subject — object acc verb Q?

(17) *pro wh*-object — verb Q?

(18) * *wh*-subject — *pro* verb Q?

Kanno's results

Kanno tested 26 college students in Japanese II on case particle drop.



Conclusion

The results are pretty strong—the L2'ers do indeed seem to be distinguishing case marker omission in the way the ECP would predict, even though the case was very convincing that they could not have been taught this.

Missing controls

There are a couple of problems with this study, however.

One is that we're comparing the naturalness of an overt object marker on a non-*wh*-word to the naturalness of a missing marker on a *wh*-word. What we don't have is any baseline for the naturalness of a case marker on a *wh*-word.

Another important problem is that there was no test of *wh*-words with a topic marker. We know that L1 Japanese speakers reject these, but it was important for the L2'ers that they *know* this, so that they can know that leaving a subject *wh*-word unmarked is omitting a nominative case marker (and not omitting a topic marker).

A more complete set of conditions would be

(19) subject wa *wh*-object — verb Q?

(20) * *wh*-subject — object acc verb Q?

(21) *pro wh*-object — verb Q?

(22) * *wh*-subject — *pro* verb Q?

(23) * *wh*-subject wa object acc verb Q?

(24) subject wa *wh*-object acc verb Q?

Allows testing: (19) vs. (24) (preference for acc drop on *wh*-phrases?), (23) (control: is *wa* disallowed on *wh*-phrases?), (19)/(21) vs. (20)/(22) (preference for acc drop over nom drop?)

Conclusions about the ECP

Kanno's conclusion: L2 learners of Japanese have nevertheless (statistically significantly, as a group) gotten the rule about dropping subject case markers, despite the lack of evidence from the textbook, the instructor, or even English.

That is, they appear to know the ECP.

Kanno takes this to mean that L2 learners must be getting this from UG, because it wasn't taught and it didn't come from the L1.
But—are we sure that it didn't come from the L1?

"UG" in L2A

The conception of UG has undergone a number of conceptual shifts...

It has always been a means of explaining how kids uniformly and quickly reach the complex knowledge system that language (L1) is.

- **Blueprint.** The LAD uses UG as a template to acquire rules. UG leads to L1 but UG is not in any way part of L1.
- **Component.** UG is the common core of language knowledge, the specifics of L1 are stored as parameters in the lexicon. There is no L1 without UG.

We probably haven't reached UG, just L1

Kanno was actually aiming to make a claim about the participation of UG in L2A, based on the idea that the application of the ECP to *case markers* is something that couldn't come from L1 but only through a re-application of UG (blueprint).

But if applying the ECP is how English speakers know **Who did you say that left?* (rather than using the ECP to learn that **Who did you say that left?*), then that same ECP can account for the case markers without "dipping into" UG (=LAD) again.

The difficulty

It is in fact probably just impossible to get at the question in the way that Kanno was trying to do it. Anything that holds true of all languages will necessarily be part of the L1.

So, if we actually understand the ECP correctly, then it doesn't have to do with *how* it is used (in "that-trace" configurations vs. case marker omission), it's about the more abstract property of "identification" of silent elements. So, it doesn't matter that the English use of it looks different from the Japanese use of it.

Most promising route

The most promising route, really, would be to look at parameter settings. In particular, one scenario that looks like it might actually reach a conclusion about UG in L2A:

Language 1 has parameter settings +A and +B. Language 2 has parameter settings –A and –B. If we find L2'ers with an interlanguage grammar that seems to have mixed parameter settings (e.g., +A, –B), then this is neither like the target language nor like the source language. The fact that this option is open would suggest that the knowledge of what the options *are* is still available to a L2'er. We'll look at that kind of configuration shortly.