1 Parameters

1.1 Universals

Language universals

Historically, as people investigated languages and learned more about how languages vary (and don't), various typological observations were made. The most thorough early attempt at this was done by Joseph Greenberg, who formulated "universals" like these:

> **Universal 3** Languages with dominant VSO order are always prepositional.
>
> **Universal 13** If the nominal object always precedes the verb, then verb forms subordinate to the main verb also precede it.
>
> **Universal 16** In languages with dominant order VSO, an inflected auxiliary always precedes the main verb. In languages with dominant order SOV, an inflected auxiliary always follows the main verb.

1.2 Parameters and clustering

Explaining universals

Theoretical explanations can then be advanced to try to explain why they hold. We can begin to understand universals 3, 13, 16 by supposing that languages have a small number of "headedness" settings determining whether the head precedes or follows its complement. IP has the same headedness as VP, CP can be different, but matches PP.

Clusters of properties

In general, the hope of determining "parameters" of language is that by virtue of a single "setting," several properties can be simultaneously derived (explaining why only certain clusters of properties occur). This can be stipulated, initially, although the further hope is that the parameters will help us understand how the properties are connected.

In the L2A context, this also helps us determine whether L2'ers are actually acquiring new settings for parameters—the properties dependent on a single parameter should go together in the interlanguage.

2 Verb movement

2.1 Parameter and clustering

Reminder: Differences in verb raising

Languages differ in whether the verb raises (as we've seen). As a reminder: The verb raises in French, not in English.

(1)   a.   Marie regarde$_i$ [souvent [$_{VP}$ t$_i$ la télévision.]]

         b.   * Marie [souvent [$_{VP}$ regarde la télévision.]]

(2)   a.   * Mary watches$_i$ [often [$_{VP}$ t$_i$ television.]]

         b.   Mary [often [$_{VP}$ watches television.]]

(3)   a.   * The children like$_i$ [not [$_{VP}$ t$_i$ spinach.]]

         b.   The children (do) [not [$_{VP}$ like spinach.]]

Verb movement "clustering"

Verb movement in this context can be said to explain a small cluster of properties.

- SVA(O), *SAV(O) (French): V moves to T.

- *SVA(O), SAV(O) (English): V does not move to T.

Implication: If a L2'er learns that SAV(O) order is possible, it follows (if they have set this parameter in the interlanguage) that SVA(O) is impossible.

2.2 White 1991: Initial study

Participants

White 1991. Native speakers of French learning English. Grades 5 and 6, very little exposure to English prior, or outside the classroom. Entered a 5-month intensive ESL program, where their schooling was devoted entirely to ESL.

Two groups: Specific instruction on: (i) English adverb placement, or (ii) question formation. Three months in, students to a pretest on adverb placement, after which the adverb group was trained on adverbs. After teaching period, students took a test, and then another at the end of the ESL program (about 5 weeks later). Finally, the (original) 5th graders were retested a year later.

Methodology

Grammaticality judgment: Cartoon story with captions; if student thought caption was incorrect, they drew arrows to repair the word order.

Preference task: Students were given a sentence in two possible orders and asked to respon if both were good, neither was good, or only one (and which one) was good.

Manipulation task: Students were given cards with words on them and told to line them up to form a sentence; then asked if they could form another with the same cards, until they couldn't continue.

Results

Grammaticality judgment task: Adverb group went from very high acceptances for *SVAO to very low (native-speaker-like) levels at the first post-test and remained there for the second one. The question group remained high throughout.

Adverb group went from moderate use of SAV to high (nearly native-speaker-like) levels at the first post-test, and remained there for the second one. The question group remained at moderate acceptance throughout.

The effect of instruction was pretty dramatic in the first and second post-tests. Explicit instruction helped. (SVAO score, SAV score). Preference task revealed the same thing.

**Results**

SVAO and SAV scores
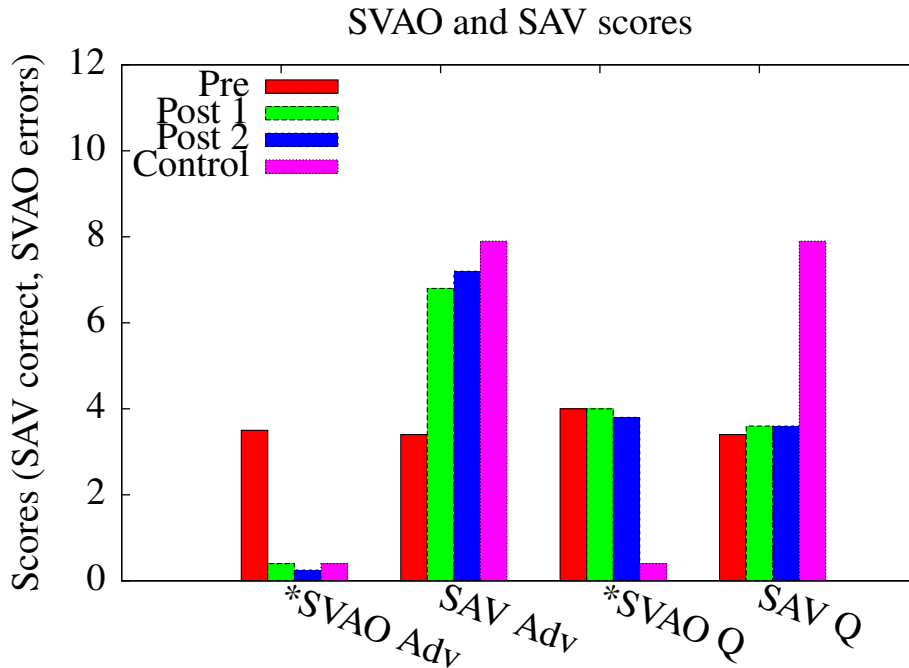
Discussion

The question group was getting basically positive evidence only (adverb position was not explicitly taught). And they didn't fare well on the tests.

The adverb group was getting explicit negative evidence and it seemed to help a lot.

...A startling result when testing those kids who were helped so dramatically by instruction: the knowledge they gained didn't last. Does this seem like a new parameter setting?

2.3   White 1991: Followup

The one-year later test

SVAO and SAV scores with followup
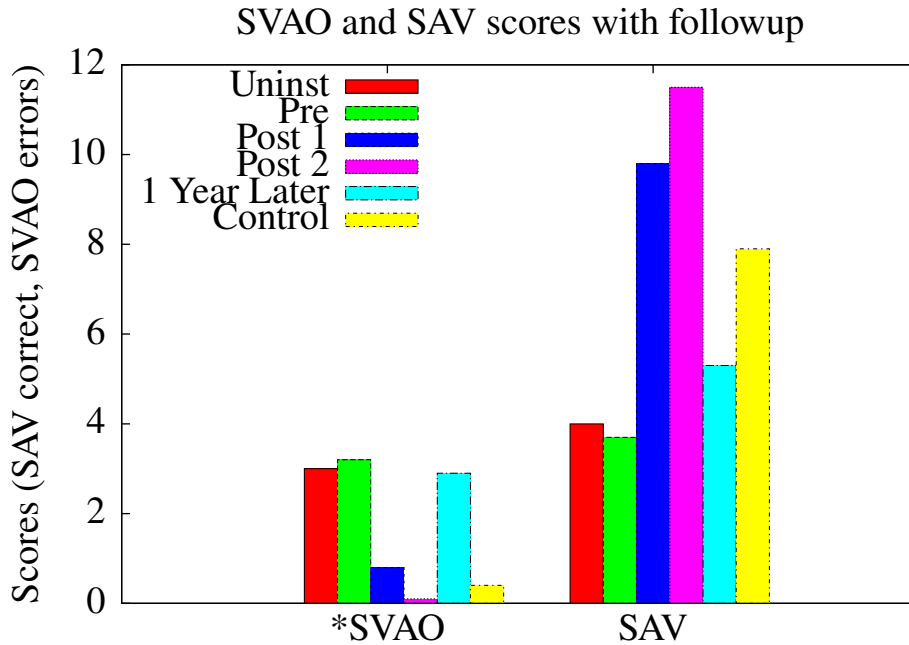
Explanations

White also observed that while her adverb group correctly ruled out *SVAO sentences in English after explicit instruction, they seemed to have incorrectly generalized this to also rule of SVAPP.

(4)   Harry runs quickly to his house.   (being rejected)

(5)   Harry quickly runs to his house.

Schwartz & Gubala-Ryzak (1992) discuss this and point out that this is not something that is possible in a natural language via parameter setting—this behavior can't be the result of mis-set parameters, it must be some kind of prescriptive rule.

2.4   Trahey 1996: Flooding

Types of input

What White (1991) was trying to test was the effects of different kinds of input—negative input via explicit instruction on adverbs vs. positive input via exposure (without concentrating on adverbs specifically). In her "positive evidence" (question) group, very little progress was made—is positive evidence ineffectual?

White speculated that the kids in the question condition might not have actually *heard* many adverbs, after listening to some tapes of the classes. Perhaps they just didn't have *enough* positive evidence?
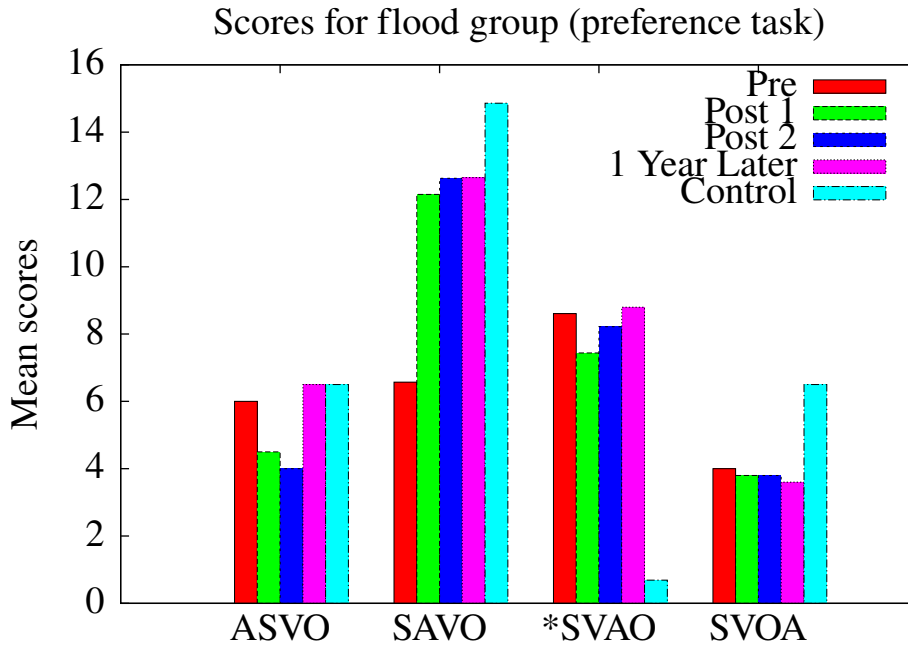
Flooding

Trahey (1996), Trahey & White (1993) set out to test this by getting together another group of students and subjecting them to an "input flood" of adverb material—no explicit teaching of adverbs, but lots of examples of proper adverb placement in English. Then they ran basically the same tests on the kids as in the other experiment, including the "one year later" test. (Trahey 1996)

The effect of the input flood appears to have been an increase in the flood group's use of SAVO, but no real change in anything else (in particular *SVAO).

Group comparison

Scores at followup
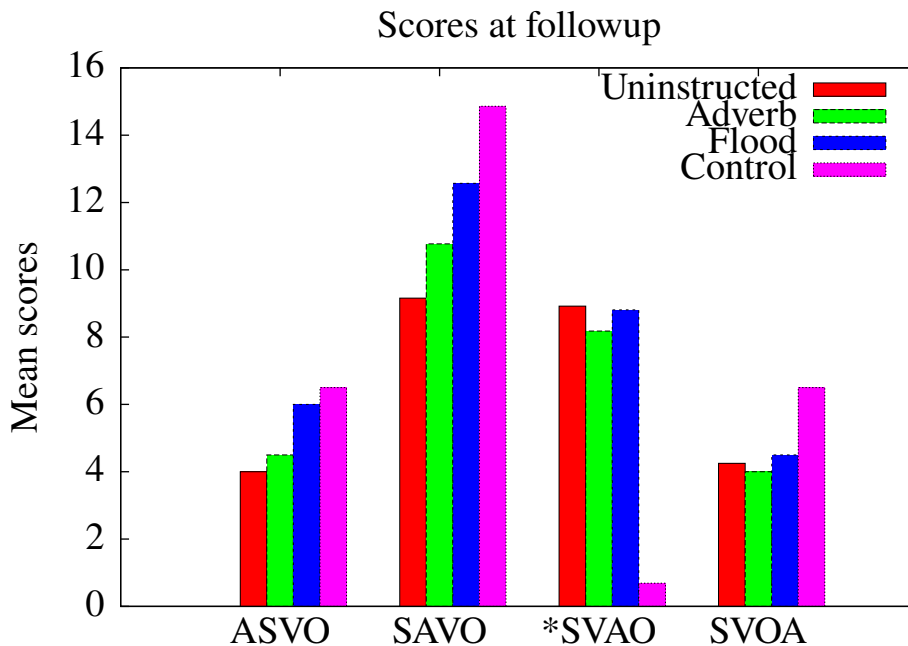
Flood group, across tests, preference task

6

Scores for flood group (preference task)

Flooding

The flooding experiment seems to have shown:

- that the knowledge gained by flooding seems to be more persistent than the knowledge gained by explicit instruction.

- that acceptance of SAVO and rejection of SVAO appear to be independent— the flooding group learned that SAVO was allowed and retained this knowledge, but still didn't reject SVAO.

- This isn't expected if the "knowledge" is a parameter setting that is supposed to have both effects.

3 Binding theory

3.1 Theoretical considerations

Advances in Binding Theory

It has been noticed that, typologically, anaphors that seem to be able to get their referent "long-distance" tend to be monomorphemic—this is particularly clear for languages that have *both* kinds of anaphors, like Dutch *zich* (LD) and *zichzelf* (local), Norwegian *seg* (LD) and *seg selv* (local), etc. (It's debatable maybe whether

this is a true correlation or a strong tendency—but for now, we'll treat it as a true correlation).

One fact about LD anaphors that seems to be pretty robust is that LD anaphors are *subject*-oriented—they can get their reference from a LD *subject*, but not from anything else outside of their clause.

Anaphor types

LD anaphors: monomorphemic, subject-oriented. Local anaphors: neither.

LD anaphors themselves fall into two types of behavior, having to do with whether they can "see out of" a finite clause or whether they can only "see out of" an infinitive clause. Finite clauses are more "opaque."

The "opacity" of finite clauses is actually a language-by-language parameter, whereas the LD/local differentiation is an anaphor-by-anaphor parameter.

Anaphor types: examples

(6) English *himself* (type 1: –LD)
   a. Fred$_i$ believes John$_i$ to have hurt himself$_{*i,j}$.
   b. Fred$_i$ believes that John$_j$ hurt himself$_{*i,j}$.

(7) Russian *sebja* 'self' (type 2: +LD–finite)
   a. Saša$_i$ poprosila Marinu$_j$ narisovat' sebja$_{i,j}$.
       'Sasha$_i$ asked Marina$_j$ to draw self$_{i,j}$.'
   b. Saša$_i$ prosit, čtoby Marina$_j$ narisovala sebja$_{*i,j}$.
       'Sasha$_i$ requests that Marina$_j$ draw self$_{*i,j}$.'

(8) Japanese *zibun* 'self' (type 3: +LD)
   a. Alice$_i$ wa Sue$_j$ ga zibun$_{i,j}$ o aisiteiru to omotteiru.
       'Alice$_i$ thinks that Sue$_j$ loves self$_{i,j}$.'

3.2 MacLaughlin 1998

MacLaughlin 1998

Suppose that English is "+opaque" with polymorphemic, local anaphors (*himself*). And Japanese is "–opaque" with monomorphemic, LD anaphors (*zibun*). That can be considered to be opposite settings on two parameters. If acquiring

English from Japanese involves setting these parameters, there are two to set—and it's at least in principle possible that one gets set before the other.

If we see Japanese ESL'ers with a "–opaque/polymorphemic" system (having set the anaphor parameter but not the opacity parameter), this amounts to the Russian system. And it is a system that differs from *both* the L1 *and* the L2 (so not derivable from evidence from either language), yet it is a possible system allowed by the parameters. This is a strong kind of evidence for UG continuing to constrain the *parametric options* in L2A.

MacLaughlin 1998

MacLaughlin (1998) looked at speakers of type 3 languages (5 native speakers of Chinese, 10 native speakers of Japanese) learning English (type 1) in various settings. What she was specifically looking to do was to classify each learner as "type 1," "type 2," or "type 3" to see in particular if there are any that show up as type 2.

Illustration

There are two parameters relevant to the type that a learner is assigned to. We can see that type 2 is not a surprising place for some learners to arrive at on the way to target type 1.

	NL	T3	T2	T1	TL
Anaphor type					
Monomorphemic	+	+	+		
Polymorphemic	(+)			+	+
AGR (finite tense blocks LD relation)					
	−	−	+	+	+

Materials

Tom thinks that John hates himself.

- *Himself* can be *John*. Agree: ____ Disagree: ____

- *Himself* can be *Tom*. Agree: ____ Disagree: ____

9

Several types of sentences were tested, including sentences with embedded finite clauses and embedded infinitival clauses with both subjects and non-subjects as potential antecedents.

Results

MacLaughlin's results were these. The 80% and 100% columns concern how consistent a subject had to be to be classified as that type. (Finite: Local only? 1. LD-Obj? Other. Else: 3. Infinitive: Local only? 1. Else: 2.—this may overestimate the number of type 2.)

L1	Type I (E)		Type 2 (R)		Type 3 (J)		Other	
	80%	100%	80%	100%	80%	100%	80%	100%
E	18	16	0	1	0	0	0	1
L2	6	4	7	4	2	5	0	2
C	3	2	1	1	1	1	0	1
J	3	2	6	3	1	4	0	1

Success?

Nontrivial numbers of subjects wound up in this "intermediate stage"—so, good evidence for UG/parameters in L2A. The parameter of the anaphor and the parameter (AGR) concerning the opacity of finite tense seem to be able to be "reset" and moreover, we see the predicted intermediate point when only one but not the other has been set to the target setting.

The Type 2 learners: consider their anaphors to be monomorphemic (LD capable), but have set the AGR parameter to "opaque."

... Though: we don't have any *independent* evidence that the "Type 2'ers" take the anaphors to be monomorphemic. White (2003) notes that monomorphemic anaphors in L1s don't show person/number agreement. Do the "Type 2'ers" use *himself*, *themselves*, *herself* correctly? We would predict not, if these are *really* Type 2 learners—right?

Izumi 2007

Izumi (2007) looked at L2'ers command of agreement, and methodology using a truth-value judgment task. She ultimately found that the ESL'ers were much more accepting of LD anaphors in non-finite clauses, so might have classified

quite a few as "type 2" in MacLaughlin's sense. But Izumi also found that there was pretty much no trouble with the gender on the anaphors—it is unlikely that the L2'ers thought *himself* was monomorphemic. Although, it's also possible that the monomorphemic/polymorphemic correlation with LD/subject orientation is not perfect. Several things still a bit up in the air.

4   Models (finishing from last time)

4.1   Full Transfer (Full Access)

Support for FTFA

The effect of the L1 on the L2 (the transfer part): Yuan 1998 (reflexives), Haznedar 1997, Vainikka & Young-Scholten 1996 (word order), Slabakova 2000 (telicity).

The other aspect of FTFA concerns whether L2'ers can "reset" the parameters to the L2 values (which we'll take up more next time), or even to values of neither the L1 nor L2 (which we'll take up more the time after that).

Falsifiability

What kinds of evidence would lead us to conclude that FTFA is *incorrect*?

White (2003) highlights that a potential problem (with FTFA as with any proposal) is that once might be able to "explain away" counterexamples (like cases where speakers of different L1s behave the same way in the L2) as simply being beyond the transfer stage.

For example, Yuan's (2001) evidence that even French L1 speakers don't raise the verb in L2 Chinese—because they've already had enough exposure to have acquired that?

4.2   Minimal Trees

Minimal Trees: the idea

Vainikka & Young-Scholten, in a series of papers, proposed the **Minimal Trees** model for second language acquisition.

The basic idea is that the starting point for second language syntax is a very reduced syntactic structure, which gets more complex over time. It is much like the Small Clause model in L1 acquisition—beginning L2'ers have syntactic structures that consist only of a VP, and as they advance, their trees become taller.

L2A takes place in stages, with grammars that successively replace each other.

Minimal Trees: Initial state and transfer
V&YS propose a certain kind of "full transfer"—but limited to the VP.

Since the initial grammar only generates VP, only parameters that affect the VP level are transferred from the L1. Most relevantly: headedness transfers.

Other parameters (such as whether the verb raises to I) do not transfer.

Minimal Trees: evidence
We saw evidence of headedness transfer (VP), but the other part of the proposal is that functional categories are missing—we're looking for the same sort of evidence we sought for in the Small Clause model of L1 acquisition.

Things associated with missing parts of the structure should be missing (or maybe default). Working backwards, if there is no C, we should expect no complementizers (*that*, *if*) and no *wh*-questions. If there is no I, we should expect no modals/auxiliaries, verb raising, or subject agreement. (Perhaps this could be made more refined by considering TP and AgrP separately.)

VP stage: data
At the VP stage, V&YS find a lack of: verb raising, auxiliaries and modals, agreement, complementizers, *wh*-movement, questions, embedded clauses. Differentiation between VP-i and VP-ii has to do with whether the head is initial (VP-i) or final (VP-ii). (All of the auxiliaries and modals came from Rosalinda (Sp.): three *wolle* 'want' and five *is(t)* 'is'. She doesn't *control* IP yet?)

stage	L1	Aux	Modal	default agr
VP	Kor	1	1	68%
VP	Tur	0	1	75%
VP-i	It	0	0	65%
VP-ii	It	0	0	82%
VP-i	Sp	8	5	74%
VP-ii	Sp	1	1	57%

TP stage: data

A little further along, some auxiliaries and modals, Korean/Turkish speakers raise the verb about 46% of the time (but note: TP in German is head-final, yet in L2 TP stage it must be assumed to be head-initial), still a lot of default agreement.

stage	L1	Aux	Modal	default agr
TP	Sp	21	9	41%
TP	Tur	0	5	68–75%

AgrP stage: Korean/Turkish speakers raising the verb 76% of the time, some embedded clauses with complementizers, complex *wh*-questions attested.

Minimal trees: assessment

The stages are not very clean—why are there *any* complementizers in the AgrP stage? Perhaps a better way to think about it is in terms of competition between AgrP and CP grammars, where the CP grammar initially loses most of the time, but gains power.

Though, also, there are NegPs and DPs, even in the VP stage, which are functional categories. And there is evidence that, e.g., English children learning French seem to manage to raise the verb. And we need to assume that some of the CP functions can be "emulated" in lower phrases (*wh*-questions in pre-CP stages, head-initial TP in order to get V2 in pre-CP stages), though again maybe this can be answered in terms of grammar competition.

4.3 Valueless features

Valueless Features

The Valueless Features hypothesis (Eubank 1993/1994) supposes that parameters in the initial state are initially "unset" (which is taken to imply variability between "on" and "off" values).

There is certainly a fair amount of variability, but there are still a number of case where the "on" setting doesn't seem to be in evidence. The primary example White (2003) points to is verb-raising—we don't see verbs raising past negation, even if they raise past adverbs, and we don't see verb raising at all in Yuan's (2001) L2 Chinese study. White (2003) also points to a number of methodological problems in the studies that even seem to support the Valueless Features hypothesis.

4.4 Conclusion

Where we are

Ultimately, it seems like something like the Full Transfer/Full Access hypothesis is closest to being able to explain what we're seeing, although we have not spent much time looking at the "full access" part of this.

Generally, there seems to be a strong effect of the L1, and there seems to be more knowledge pertaining to the higher functional structure in the interlanguage than would be expected on the Minimal Trees hypothesis, and more constraint than would be expected on the Valueless Features hypothesis.