OT and predicting how optional infinitives really are
GRS LX 865, 10/18/04

Optimality Theory

Optimality Theory, although a general system, arose out of work in phonology.

Phonology was traditionally done in a rule-based fashion, with intermediate representations, much like the syntax we are familiar with. There is a starting point (*underlying form*), there are a set of rules which apply (perhaps in a particular order, perhaps over and over again) until the *surface form* is reached.

But it was observed that there are phenomena which, although easy enough to explain in "plain English," came out strangely or only in a roundabout way from the rules.

Let's do a little bit of phonology. We'll look at the phonology of Mohawk, which can illustrate this reasonably well.

(1) Penultimate stress (normally)

a. r á k . w a s
   2 1

a. rákwas /hra-kw-as/ 'He picks it'
b. wakashé:tu /wak-ashet-u/ 'I have counted it'
c. wakharatatuhátye /wak-haratat-u-hatye/ 'I go along lifting up'

(2) Contexts in which epenthetic *e* is inserted:
a. Between a consonant and a single sonorant (C*e*n, C*e*r, C*e*w)
b. Between a consonant and a word-final glottal stop.
c. After a consonant when followed by a consonant cluster (except *h*C and *s*C).

(3) In some cases, stress skips epenthetic vowels (*ɐ* below)

a. t é k ⓔ r i k s
   2 1

a. tékɐriks /te-k-rik-s/ 'I put them together'
b. wákɐras /w-akra-s/ 'it smells'
c. tákɐrikɐ? /t-ʌ-k-rik-ʔ/ 'I'll put together side by side'
d. waʔtkatátɐnakɐ? /waʔ-t-k-atat-nak-ʔ/ 'I scratched myself'

The way this has traditionally been handled is with a set of rules that happen in an order.

(4) ASSIGN STRESS (to the penultimate syllable)
INSERT EPENTHETIC VOWELS

But *not all* epenthetic vowels are skipped for stress:

(5) Unskipped epenthetic vowels (*e* below)

a. w a k ⓔ n y a k s
   2 1

a. wakényaks /wak-nyak-s/ 'I get married'
b. tekahsutérhaʔ /te-k-ahsutr-haʔ/ 'I splice it'
c. sasáhket /sa-s-ahkt/ 'go back!'

The natural change to make is to suppose that there are some epenthetic vowels inserted *before* stress assignment.

(6) INSERT SOME EPENTHETIC VOWELS
ASSIGN STRESS
INSERT SOME MORE EPENTHETIC VOWELS

Here's Michelson's (1989) proposal about Mohawk stress and epenthesis:

(7) Joiner Insertion (insert *a* to melody tier in C]_[C)
e-Epenthesis (insert *e* to melody tier in C_R, C_?#)
V-Insertion I (link unlinked vowel to V slot in C_CC)
Mohawk Stress Rule (stress penult)
V-Insertion II (link unlinked *a* [joiner] to V slot)
Vowel Lengthening (lengthen stressed vowel in _CV)
Prothesis (insert *i* in #_C_0VC_0#)
V-Insertion III (link unlinked *e* to V slot)

What this says is:
Insert an (invisible) epenthetic *e* where called for.
If any of them are before a consonant cluster, make it visible.
Assign stress.
Any invisible *e* that's left over, make it visible.

But there's another way to look at this:

Epenthetic *e* does not count for stress assignment *unless it is needed for syllabification*.

A proper syllable in Mohawk is a CVC syllable. Go back and look at where *e* counts. We can say that Mohawk has a constraint against counting epenthetic *e* for stress, but that it is overridden by a constraint requiring syllables to have the shape CVC (CVC*s* counts as CVC). Isn't that much easier?

(8)

/te-k-rik-s/	CVC	*STRESS-E
☞ a. ték<e>riks		
b. te.ké.riks		*!

(9)

/wak-nyak-s/	CVC	*STRESS-E
a. wák<e>nyaks	*!	
☞ b. wa.kén.yaks		⊛

The structure of the theory

(10) Constraints are universal (all lgs have them—*this* is UG)
Grammars differ only in the ranking of the constraints.
The child's task during acquisition is to acquire the *rankings*.

		EVAL C_1	C_2 …	
	candidate$_1$	*	* …	
	candidate$_2$	*	✔	
	candidate$_3$	✔	✔	
/underlying / → **GEN** →	candidate$_4$	✔	*	→ ☞ candidate$_3$
	candidate$_5$	✔	*	
	candidate$_6$	*	✔	
	candidate$_7$	✔	✔	
	…			

Optimality Theory has pretty much "taken over" phonology—but it hasn't gotten very far in syntax yet, although there are a number of people working on a number of different OT-based approaches…

A problem with an optimality character: Expletive subjects in German

(11) a. Es wurde schön gedanzt.
it was beautifully danced
'The dancing was beautiful.'

b. Schön wurde getanzt.

c. * Schön wurde es getanzt.

The question: Why is (11c) ungrammatical?
Seems: *es* can satisfy the EPP *if nothing else does*.

A simple case:

(12) a. It rained. (English)
b. Piove. (Italian)

We could say that in Italian the EPP is satisfied, but by *pro*—but there is also a sense in which this renders the EPP as a universal (inviolable) constraint pretty vacuous.

Plus, there is something *better* about (12b)—it has no meaningless elements in it.

Suppose that we have two constraints:

(13) SUBJECT The highest A-specifier must be filled. (EPP)
FULL-INT Lexical items must contribute to the interpretation.

For sentences that mean 'it rained' you can't have it both ways. There is no argument, so if you are going to have the subject position filled (satisfying SUBJECT) you have to fill it with a meaningless element (violating FULL-INT).

In English, it is more important to satisfy SUBJECT, even if that means violating FULL-INT.

(14) English: SUBJECT >> FULL-INT

In Italian, it is more important to satisfy FULL-INT, even if that means violating SUBJECT.

(15) Italian: FULL-INT >> SUBJECT

The idea is that both SUBJECT and FULL-INT are universal constraints on language, but languages just differ as to how important each is relative to the other.

As for acquisition

We will explore the idea that a child's task is to rerank constraints to reach the ranking that the adult has.

There are a couple of possibilities.

Tesar & Smolensky (1993) **Constraint Demotion**
Evaluates constraint into strata based on matches between input and output, works its way to the eventual ranking.

Legendre et al. (2000, etc.) **Partial rankings**
Rankings of individual constraints are allowed to "float" over a range of the constraint hierarchy—we will review this shortly.
Also a possible account of optionality in adult language.

Boersma, Hayes, et al. **Gradual Learning Algorithm**
Individual constraints are stochastically evaluated—that is, there is a certain degree of "noise" in the evaluation of a location of a constraint, resulting in a normal distribution around a mean. Occasionally, this can cause a higher-ranked constraint to lose out to a lower-ranked constraint. This accounts for optionality in adult grammar, moving the means gradually to match the input accounts for child acquisition.

1. The "optional infinitive stage"

(16) **The Optional Infinitive Stage** (Wexler 1998)
a. Root infinitives are possible grammatical sentences for children in this stage (around 2 years)
b. These infinitives coexist with finite forms
c. The children nevertheless know the relevant grammatical Principles, e.g. head movement, checking, etc.

An example of this from French:

(17) ‡ Cabinets ouvrir.
(Grégoire 1;9.28)
Restroom open-INF
'(I will) open the restroom (door)'

Verb in (17) is the **infinitive** *ouvrir* (vs. the correct first-person form *ouvre* [present])

Evidence has been found for an OI stage of this sort in at least: Danish, English, Faroese, Icelandic, Norwegian, Swedish, French, Irish, Brazilian Portuguese, Czech. (see Wexler 1998 for citations)

2. The French data: Children and stage divisions

French child data came from the CHILDES database (MacWhinney & Snow 1985): Grégoire (between 1;9 and 2;5), Stéphane (between 2;2 and 3;3), and Philippe (between 2;1 and 2;6).

PLU stages (Predominant Length of Utterance) grouped kids at similar levels of development together more reliably than MLU (Mean Length of Utterance) or age (Vainikka, Legendre, and Todorova 1999).

Table 1. Children, files, and ages included in this study

a. Grégoire (Champaud corpus)

Files	Age	PLU stage	Total # of Utterances
1–4	1;9–1;10	3b	874
5–7	2;0–2;3	4b	732
8–10	2;5	4c	1038

b. Stéphane (Rondal 1985; Rondal, Bachelet, and Leveillé 1973)

Files	Age	PLU stage	Total # of Utterances
1–3	2;2–2;3	3b	644
6a/6f/8a	2;6–2;8	4b	688
25b	3;3	4c	257

c. Philippe (Suppes, Smith & Leveillé 1973)

Files	Age	PLU stage	Total # of Utterances
1–3	2;1–2;2	4b	898
11	2;6	4c	387

(18) *PLU stages in our data*:

Stage 3: *"Two-word" stage*
- The one-word sentence type no longer clearly predominates (i.e. fewer than 60% of all utterances are one-word utterances)
- Of the three sentence types, the multi-word sentence type is not the most common one

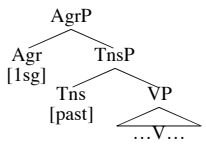
Stage 4: *Predominantly multi-word stage*
- Of the three sentence types, the multi-word sentence type is the most common one

(19) *Secondary PLU stages in our data*
Secondary stage b: 11%–60% of utterances contain a verb
Secondary stage c: over 60% of utterances contain a verb

3. Representation of tense and agreement

Adult 1st past tense form has the following (partial) structure:

(20)



It has been proposed that in certain child utterances, one or both of AgrP and TnsP can be missing from the representation (Schütze & Wexler 1996, Wexler 1998, Ingham 1998).

- If *both* Agr and Tns are missing, a **nonfinite form** results.
- If Agr is missing, **default agreement morphology** surfaces.
- If Tns is missing, **default tense morphology** surfaces.
- If both Agr and Tns are there, **tense and agreement morphology** surface.

Well-known: (crosslinguistically, but even in French specifically) first agreement morphology to appear is **3sg**, and the first tense morphology is **present** tense. For a time, they will **overgeneralize** 3sg or present tense, using it where non-3sg or non-present tense would have been appropriate in the adult language.

(21) Papa et Maman est parti. (Grégoire 2;0.5)
Father and Mother is gone
'Mother and Father have gone.'

We thus have **four possibilities** for verb forms in the child data:

• Tensed and agreeing	*j'ai dansé*	1sg, past	**finite**
• Tensed, not agreeing	*a dansé*	3sg, past	
• Agreeing, not tensed	*je danse*	1sg, present	
• Neither tensed nor agreeing	*danser*	NRF	**nonfinite**

A note about subject agreement: We assume subject clitics (*je*, *tu*, *il*, …) are overt realizations of agreement—specifically, we do *not* take these to be overt subjects. This means that ***je danse*** '(I) dance' is analyzed as a sentence with a **null subject** and **overt 1sg agreement**. On this point, we follow Auger (1994, 1995), Cummins & Roberge (1993), Ferdinand (1996), Hulk (1986), Kaiser (1994), Lambrecht (1981), Legendre (1999), Paradis & Crago (1999), Pierce (1992), Roberge (1990), Suñer (1988), and others.

4. The French data: What we found

Because **3sg is the default agreement**, and **present is the default tense**, we counted only **non-3sg** (unambiguous agreement) and **non-present** (unambiguous tense).

Downside: sometimes the speaker *intends* to use a 3sg subject or present tense.

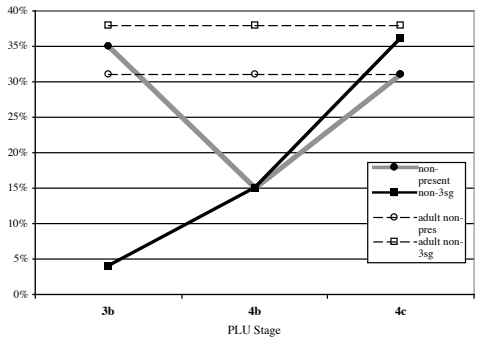
Table 2: Adult usage of non-3sg and non-present tense

Adults from file	non-present	non-3sg
Grégoire 9	28% (184/661)	35% (231/659)
Philippe 11	34% (173/507)	41% (206/506)
Average	31% (357/1168)	38% (437/1165)

- If a kid uses non-present about 31% of the time, s/he's representing tense at 100%.

- If a kid uses non-3sg about 38% of the time, s/he's representing agreement at 100%.

Figure 1. Tense and Agreement (tables 2, 3, 4)

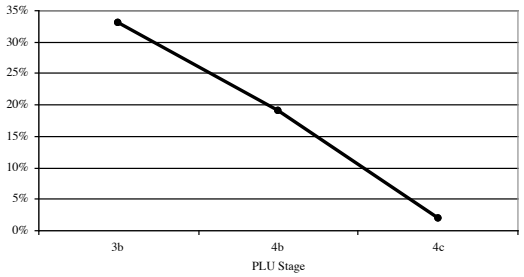


Pattern: Tense starts out being used at adult-like levels (St. 3b).
Tense use drops dramatically (Stage 4b).
Tense returns to adult-like levels (Stage 4c).

Agreement starts out essentially unused (Stage 3b).
Agreement increases (Stage 4b).
Agreement reaches adult-like levels (Stage 4c).

☞ **Tense and agreement show distinct courses of acquisition.**

Figure 2. Proportion of non-finite root forms of all verbs (table 5)



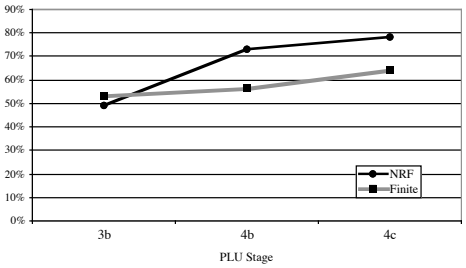
5. Constraining analyses: Null subjects

It has been long observed that children at this age tend to show a great many "null subjects". These occur both with finite and nonfinite verbs.

> **Note**: Recall that we take subject clitics to be realizations of agreement, meaning that *je partis* 'I leave', grammatical in the adult language, would constitute a finite null subject utterance. To put it another way, **adult French *allows* null subjects** from this perspective.

It turns out that the rate of null subjects in NRFs is somewhat higher than in finite verbs, suggesting that there is something particularly amenable to null subjects in the NRF structure. (This won't follow from anything here).

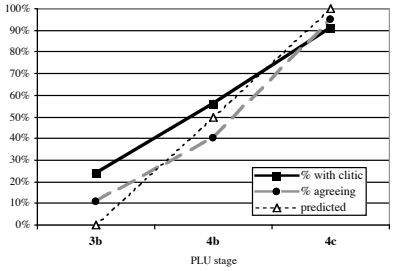
Figure 3. Null subjects (tables 6, 7)



There are **two kinds of finite null subjects**:
those with agreement (*a subject clitic*)
those without agreement.

If the ability of a finite verb depended on agreement (either being present or being absent), we would expect to see that the proportion of agreement among finite null subjects would be different from what we find in the pool of verbs as a whole.

Figure 4. Finite null subjects with agreement clitics (table 8)



What this means: The rate of agreement in finite verbs when restricted to null subject cases is the same (basically) as when unrestricted. ***Conclusion***: Whatever lets you drop the subject doesn't depend on (but rather is orthogonal to) the presence of agreement features.

6. Constraining analysis: postverbal subjects

It has been observed that kids around this age will produce sentences like (22a), with a postverbal subject. Such sentences are ungrammatical in adult French. These kids also produce sentences like (22b), a "right dislocation", grammatical in adult French.

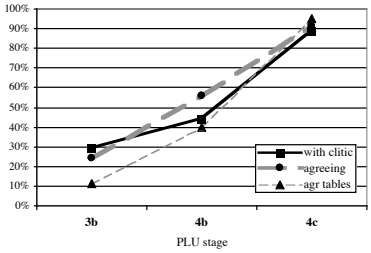
(22) a. ‡ Manger salade Adrien (Grégoire 1;9.10)
eat-INF salad Adrien
'Adrien is eating salad,' *(the clitic is the difference)*

b. Il est monté Grégoire (Grégoire 1;9.28)
he has gone.up Grégoire
'Grégoire has gone on top.'

Some (e.g., Déprez & Pierce 1993), have argued that (22a) is a special kind of error, in which the subject gets "left behind" (in SpecVP, under the VP-Internal Subject Hypothesis). Others (e.g., Ferdinand 1996, Labelle & Valois 1996) argue that (22a) is just a variant of grammatical (22b) but missing the clitic.

On our view, the clitic in (22b) is simply **agreement**. Therefore, if (22a) is just (22b) without agreement, we would expect that the proportion of (22b)-type to (22a)-type postverbal subjects would be the same as the overall rate of agreement.

Figure 5. Proportion of postverbal subjects with agreement clitic (table 9)



What this means: Like with the null subject case considered before, this tells us (*a*) that the driving force behind Right Dislocation (of the (22b) type) cannot depend on the presence or absence of agreement, and (*b*) that (22a–b) are instances of the same construction, differing only in the presence/absence of agreement features.

Wrap-up: Although we do not have an analysis of either the null subjects or the postverbal subjects in child/adult French, the data we have has limited the types of analysis that are available (*This is a good thing*). Neither process can crucially rely on the presence or absence of agreement features, and the child postverbal subjects (of type (22a)) should be analyzed as the same construction as child/adult right dislocations (of type (22b)).

7. The Agr/Tns Omission Model

Why do children show "optional infinitives" in the first place?

(23) **Agr/Tns Omission Model (ATOM)** (Wexler 1998, Schütze & Wexler 1996)
   a. Agr or Tns (or both) may be deleted
   b. Morphology is inserted according to the Elsewhere principle (morphology has defaults)
   c. Kid knows adult syntax and morphology (features, Elsewhere, defaults)

• Children's utterances are *grammatical* even by adult standards.
   • The requirement that Tns be in the clause is *not* a syntactic principle. (rather, it is pragmatic, "anchoring" the interpretation to the context).
   • Similarly for Agr; if Agr is omitted, no *syntactic* principles are violated.

• Something adults know requires Agr and Tns (e.g., the "anchoring" constraints). Something that says "REALIZE AGREEMENT" and "REALIZE TENSE."

• Children in the Optional Infinitive stage have an additional constraint, the **Unique Checking Constraint** (UCC):

   (24) **Unique Checking Constraint (UCC)** (Wexler 1998) The D-feature of DP can only check against one functional category.

• (In certain types of languages) **the UCC prohibits representations with *both* AgrSP and TnsP**. *The idea*: the subject must move to "check" a "D-feature" (determiner category feature) in the specifier of each functional category. The UCC prohibits doing this twice, but if both AgrSP and TnsP are present, each will require the movement.

• **The cause of the "optional" infinitives is a decision the kid has to make about which constraint to violate.** The kid makes different choices for different utterances…

   —Violate UCC (adult utterance) (satisfying "REALIZE TENSE, AGREEMENT")
   —Violate "REALIZE TENSE" (non-tensed utterance) (satisfying UCC)
   —Violate "REALIZE AGREEMENT" (non-agreeing utterance) (satisfying UCC).

Note: Although Schütze & Wexler (1996) discuss the case where *neither* Tns nor Agr is present (suggesting that this is the source of genitive case subjects) I see no way for such a representation to come out of their system. It would mean that the child chooses to violate *both* "REALIZE TENSE" and "REALIZE AGREEMENT" even though violating only one of them would allow satisfaction of the UCC.

Another important point:

ATOM characterizes what kinds of utterances *may* occur. It does *not* characterize *how often* each occurs.

The data I presented above shows that there is a **systematic pattern** in the realization of tense and agreement features. That is, there is a system to the choices the kid makes about which constraint to violate. ATOM has nothing to say about this…

8. Analysis: The intuitive idea

There are **conflicting constraints** at play in the child

   • Sentences should show tense
   • Sentences should show agreement

but

   • Sentences should not be so complicated as show inflection.
   • Sentences should not be so complicated as to realize *both* tense & agreement. (analog to UCC)

(Whether the last two should be considered to be an attempt to "simplify" for reasons of computational complexity in underdeveloped cognitive resources is an open question, but it gives us an intuitive way to think about them).

The first group endorse **faithfulness** to the underlying representation (the "intent")

The second group endorse **minimal structure**.

The pattern we see in the kids results from the **minimal structure** constraints taking priority over the **faithfulness** constraints (thus, tense and/or agreement are not realized) initially. Then, the faithfulness constraints **become more dominant** as the child's development progresses.

9. Background: Optimality Theory

Optimality Theory (Prince & Smolensky 1993) provides a way formalize the idea

(25) (i) Grammar is an optimizing system of universal well-formedness constraints on linguistic forms.
   (ii) Well-formedness constraints are simple and general. They routinely come into conflict and are surface-violated.
   (iii) Conflicts are resolved through hierarchical rankings of constraints, which are language-particular.
   (iv) Alternative structural realizations of an input ("candidates") compete. The candidate which best satisfies (or minimally violates) the full set of ranked constraints is the optimal one. Only the optimal structure is grammatical. Every competition yields an optimal output.
   (v) Candidates are evaluated against a *strictly ranked* set of constraints; for every two constraints C_1 and C_2, either C_1 outranks C_2 or C_2 outranks C_1.

(26) **Economy of Structure constraints**
   *F: No functional heads
   *F^2: No pairs of functional heads

(27) **Faithfulness constraints**
   PARSET: Parse Tense
   PARSEA: Parse Agreement

Possible structures (candidates) evaluated for "optimalness":

(28) a. VP — *danser* (NRF); *: PARSEA, PARSET; √: *F, *F²

b. TP (T[PAST], VP) — *a dansé* (3sg, past); *: PARSEA, *F; √: PARSET, *F²

c. AgrP (Agr[1SG], VP) — *je danse* (1sg, present); *: PARSET, *F; √: PARSEA, *F²

d. AgrP (Agr[1SG], TP (T[PAST], VP)) — *j'ai dansé* (1sg, past); *: *F (twice), *F²; √: PARSEA, PARSET

Restating the intuitive idea in terms of these structures:

(29) Stage 3b: Ranking sometimes permits a single functional projection, which, when present, invariably realizes tense.

Stage 4b: Ranking permits a single functional projection; variation in ranking permits either tense or agreement to be realized.

Stage 4c: Ranking permits two functional projections, both tense and agreement are realized.

Why this won't work as-is:

In Stage 3b, realizing tense seems to be more important than avoiding functional projections altogether.

Suppose the ranking is: $*F^2$ >> PARSET >> *F >> PARSEA

Then *every verb the child tries to say* will be tensed and non-agreeing (28b).

But we know that this isn't what happens:
- 33% of the time, the verbs are not tensed *or* agreeing.
- 66% of the time, the verbs are tensed, not agreeing.

And no matter what ranking we try, we'll always end up with one of the candidates in (28) winning *all of the time*.

10. Partial constraint ranking

The solution to the problem is the ***partial ranking*** (as explored by Reynolds 1994, Anttila 1997, Boersma 1997, Nagy & Reynolds 1997).

A partial ranking can give us *proportions*. A simple example:

(30) a. Partial ranking:
Fixed: $*F^2$ >> *F
Floating: PARSET ———————

⇒

b. Set of rankings:
i. $*F^2$ >> *F >> PARSET — wins: untensed verb
ii. $*F^2$ >> PARSET >> *F — wins: tensed verb

ParseT covers a ***range*** of the ranking space.

Prior to each evaluation, it is fixed randomly in that range. So each evaluation could **either** be using (strict) ranking (30bi) **or** (30bii).

Assuming it is random, and that anyplace in PARSET's range has an equal probability of being the place where ParseT is fixed in a given evaluation:

- *untensed verb* wins 50% of the time
- *tensed verb* wins 50% of the time

So we expect to see 50% tensed verbs and 50% untensed verbs.

8. Analysis

The constraints $*F^2$ and *F are part of a power hierarchy;
$*F^2$ >> *F in every possible grammar..

The rest of the analysis relies on the distribution of PARSEA and PARSET with respect to $*F^2$ and *F.

In Stage 3, PARSET has a higher range than PARSEA—
At this stage, PARSEA is ranked so low as to be irrelevant.

(31) *Stage 3b*
Fixed: $*F^2$ >> *F
Floating: PARSET ——————————
PARSEA —

This partial ranking encodes three strict orderings.
Two yield *tensed verbs*, one yields *NRFs.*

(32) *Stage 3b*:
a. PARSET >> $*F^2$ >> *F >> PARSEA — yields: tensed
b. $*F^2$ >> PARSET >> *F >> PARSEA — yields: tensed
c. $*F^2$ >> *F >> PARSET >> PARSEA — yields: NRF

This predicts: 33% of verbs at Stage 3b will be NRFs
67% of verbs at Stage 3b will be tensed (100% of finite)
no verbs at Stage 3b will be agreeing
(no verbs at Stage 3b will be agreeing & tensed)

That's what we saw. 33% were NRFs at Stage 3b,
100% of finite verbs tensed, non-agreeing.

In Stage 4b, PARSEA advances to cover the same ground as PARSET. PARSEA and PARSET are now completely symmetrical:

(33) *Stage 4b*:
Fixed: $*F^2$ >> *F
Floating: PARSET ——————————
PARSEA ——————————

The partial ranking in (33) defines the following 12 orderings.

(34) *Stage 4b*:
a. PARSET >> PARSEA >> $*F^2$ >> *F — tensed, agreeing
b. PARSEA >> PARSET >> $*F^2$ >> *F — tensed, agreeing
c. $*F^2$ >> *F >> PARSET >> PARSEA — NRF
d. $*F^2$ >> *F >> PARSEA >> PARSET — NRF
e. $*F^2$ >> PARSET >> PARSEA >> *F — tensed
f. $*F^2$ >> PARSEA >> PARSET >> *F — agreeing
g. PARSET >> $*F^2$ >> PARSEA >> *F — tensed
h. PARSEA >> $*F^2$ >> PARSET >> *F — agreeing
i. PARSET >> $*F^2$ >> *F >> PARSEA — tensed
j. PARSEA >> $*F^2$ >> *F >> PARSET — agreeing
k. $*F^2$ >> PARSET >> *F >> PARSEA — tensed
l. $*F^2$ >> PARSEA >> *F >> PARSET — agreeing

Predicts:
2/12 = 17% yield tense & agreement
2/12 = 17% yield NRFs
4/12 = 33% yield tensed (non-agreeing) (6/10=60% of fin.tensed)
4/12 = 33% yield agreeing (non-tensed) (6/10=60% of fin.agreeing)

Of finite, we expect 60% to be tensed, 60% to be agreeing.
Of tensed, we expect 31% to be non-present, 38% to be non-3sg.
So we expect 19% (60% × 31%) of finite verbs to be non-present
And we expect 23% (60% × 38%) of finite verbs to be non-3sg.

And that's (pretty close to) what we saw.
19% were NRFs at Stage 4b. (vs. 17%)
15% of finite verbs showed non-present. (vs. 19%)
15% of finite verbs showed non-3sg (vs. 23%)

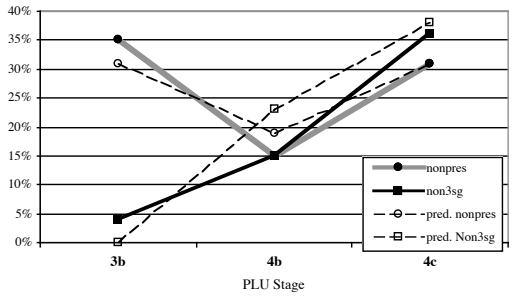
Finally, at Stage 4c, PARSET & PARSEA, together, come to outrank $*F^2$, at which point tense and agreement are always both realized.

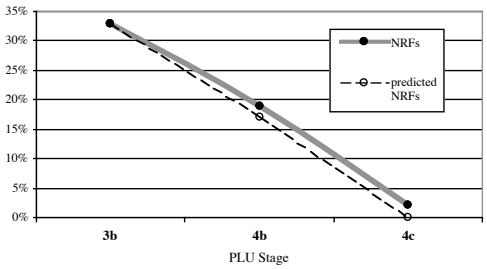
(35) *Stage 4c*:
Fixed $*F^2$ >> *F
Floating: PARSET — —
PARSEA — —

(36) *Stage 4c*:
a. PARSET >> PARSEA >> $*F^2$ >> *F tensed, agreeing
b. PARSEA >> PARSET >> $*F^2$ >> *F tensed, agreeing

Predicts: 0% NRFs, 100% tensed (31% non-present), 100% agreeing (38% non-3sg)
We see: 2% NRFs, (100% tensed) 31% non-present, (95% agreeing) 36% non-3sg.

Figures 6–7.Predictions of the partial ranking analysis vs. observed data





11. $*F^2$ vs. the UCC.

Is UCC really $*F^2$?

Well, no, but it's close. The difference between the two is that the UCC is a fairly specific constraint, whereas $*F^2$ is pretty general. UCC pinpoints a particular *reason* for disallowing two functional projections (namely, each would require a D-motivated movement). From this, Wexler (1998) can extend the analysis to capture the NS/OI generalization:

(37) **The Null-Subject/OI Generalization (NS/OI)** (Wexler 1998)
Children in a language go through an OI stage if and only if the language is *not* an INFL-licensed null-subject language.

(38) **Axiom for the NS/OI to follow from the UCC** (Wexler 1998)
In null-subject languages, Agr is D; in non-null-subject languages Agr is not D (rather, it *needs* D). Only Agr which *needs* D drives a D-motivated movement.

One problem with the NS/OI generalization is that, looking at French the way we have, it does not hold of French (under our assumptions, French *is* in fact a null subject language, yet it clearly has an OI stage). Do we *want* to derive it? (This remains an open question—there seems to be a difference in behavior between languages that "have an OI stage" and those which don't, and it may well still have to do with the idea in the axiom, that Agr in some languages does not need D. What needs to be clarified is the meaning of "null subject language" in this context in light of the analysis of French here).

12. What we've seen…

We have seen new results from child French that show:
- Realization of tense and agreement proceeds independently.
- Specifically:
  - —tense starts (at the point at which we start observing) at an adult-like level, then *dips* (in response to competition from agreement), and finally regains adult-like levels.
  - —agreement starts essentially completely unrepresented and increases roughly linearly across the three stages.
  - —NRFs start out at high levels (roughly 1/3) and decrease roughly linearly to zero.

This data also suggests that:

- Availability of null subjects is independent of realization of agreement.
- Availability of right dislocation is independent of realization of agreement.
- child postverbal subject "errors" are a form of right-dislocation.

Finally, I presented an analysis of the course of development in terms of a partial ranking of constraints in an Optimality-Theoretic framework that predicts the relative frequencies of the different sentence types (Tns/Agr, Tns, Agr, NRF), which complements the (predictively successful) ATOM model of child language proposed by Schütze & Wexler (1996), Wexler (1998).

Boersma & Hayes, Praat, and the GLA

Praat, as it happens, was written by Paul Boersma, and has a built-in version of his GLA, so we can explore a bit about how that system learns rankings.

The first step will be to look at Finnish adult data. Arto Anttila collected a corpus of Finnish genitive plurals, which it turns out speakers do not always produce the same way.

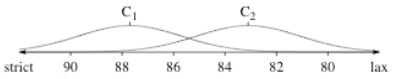
korjaamo 'repair shop' sometimes comes out in the genitive plural as **kór.jaa.mo.jen** (80%) and sometimes as **kór.jaa.mòi.den** (20%). Anttila put together a model of partial rankings that modeled this variation and predicted percentages, similar to (but not identical) to the one we've discussed for French above.

Boersma provides this as a test case, to show that the GLA can learn even a distribution of forms with variation in them. So, we'll try teaching a "kid" the Finnish genitive plural using the GLA to see how Praat works.

What we'll do in the lab is to take the French data used above and subject it to the GLA. The GLA will make a mess of it (just to give away the ending), and the last task will be to try to understand why.

Boersma's (and the GLA's) assumption about how constraints are evaluated looks like this: C_1 outranks C_2, but they are each distributed normally around a mean ($\mu 1 > \mu 2$). In a certain (computable) percentage of the evaluations, C_2 will actually win.

(6) *Overlapping ranking distributions*



Appendix 1. Tables of actual numbers

Table 3: Verbs with non-present tense inflection (out of unambiguously tensed verbs)

Child	Stage 3b	Stage 4b	Stage 4c
Grégoire	34% (66/194)	21% (44/212)	32% (205/646)
Stéphane	37% (19/52)	10% (17/179)	25% (34/135)
Philippe		13% (44/334)	30% (74/246)
Weighted avg.	35% (85/246) **113%**	15% (105/725) **48%**	31% (313/1027) **100%**

Table 4: Verbs with non-3sg agreement inflection (out of unambiguously agreeing verbs)

Child	Stage 3b	Stage 4b	Stage 4c
Grégoire	3% (5/156)	19% (33/172)	34% (221/650)
Stéphane	5% (2/43)	12% (13/109)	38% (51/133)
Philippe		15% (44/303)	40% (98/246)
Weighted avg.	4% (7/199) **11%**	15% (90/584) **40%**	36% (370/1029) **95%**

Table 5: Proportion of non-finite root forms (NRFs) of all verbs

Child	Stage 3b	Stage 4b	Stage 4c
Grégoire	28% (83/297)	18% (51/287)	1% (7/711)
Stéphane	48% (51/106)	13% (27/205)	2% (3/152)
Philippe		22% (105/476)	6% (14/250)
Weighted average	33% (134/403)	19% (183/968)	2% (24/1113)

Table 6: Proportion of NRFs appearing with null subjects

Child	Stage 3b	Stage 4b	Stage 4c
Grégoire	46% (38/83)	78% (35/45)	100% (5/5)
Stéphane	53% (27/51)	54% (13/24)	0% (0/3)
Philippe		75% (76/102)	90% (9/10)
Weighted average	49% (65/134)	73% (124/171)	78% (14/18)

Table 7: Proportion of finite verbs appearing with null subjects (clitic subjects ⇒ null)

Child	Stage 3b	Stage 4b	Stage 4c
Grégoire	53% (100/189)	60% (120/200)	60% (361/605)
Philippe		53% (143/271)	79% (150/190)
Weighted average	53% (100/189)	56% (263/471)	64% (511/795)

Table 8:Proportion of finite null subject sentences with agreement clitic

	Stage 3b	Stage 4b	Stage 4c
verb alone	76	117	44
verb with clitic	24	146	467
% with clitic	24%	56%	91%
adjusted % agr.	11%	40%	95%

This line we got from Table 4 (observed non-3sg), here expressed as a percentage of *adult* rate of non-3sg in order to give an estimate of *total* rate of agreement (i.e. including 3sg).

Table 9:Proportion of postverbal subjects with agreement clitic

	Stage 3b	Stage 4b	Stage 4c
verb alone	50	70	7
verb with clitic	20	54	59
% with clitic	29%	44%	89%
agreeing (Table 8)	24%	56%	91%

This line we got from Table 8 (proportion of finite null subjects with an agreement clitic). We use this, rather than Table 4 because it does not rely on the somewhat unreliable estimate of *adult* production of non-3sg.

Agreement in French. Consider the most common conjugation class in French, verbs ending in *-er*, exemplified in (39). Homophony is rampant (although spelling differentiates "agreement" on the verb).

(39) *The -er conjugation (French)*

		singular	phonetic		plural	phonetic
1.	(je)	**dans**e	[dās]	(nous)	**dansons**	[dāsō]
2.	(tu)	**dans**es	[dās]	(vous)	**dansez**	[dāse]
3.	(il, elle, on)	**dans**e	[dās]	(ils, elles)	**dans**ent	[dās]

Subject clitics tell us which form the verb is in—we take as the *realization of agreement*.

Tense in French.

(40) *Passé composé* 'danced'

	singular	phonetic	plural	phonetic
1.	j' **ai dansé**	[e dāse]	nous **avons dansé**	[avō dāse]
2.	tu **as dansé**	[a dāse]	vous **avez dansé**	[ave dāse]
3.	il **a dansé**	[a dāse]	ils **ont dansé**	[ō dāse]

(41) *Future Proche* 'going to dance'

	singular	phonetic	plural	phonetic
1.	je **vais dans**er	[ve dāse]	nous **allons dans**er	[alō dāse]
2.	tu **vas dans**er	[va dāse]	vous **allez dans**er	[ale dāse]
3.	elle **va dans**er	[va dāse]	elles **vont dans**er	[vō dāse]

We only counted **both the auxiliary and the participle** as **tensed** (kids drop aux).

References (some)

Anttila, Arto (1997). *Variation in Finnish phonology and morphology*. Ph.D. dissertation, Stanford University.

Ferdinand, A. (1996). *The development of functional categories: The acquisition of the subject in French*. Dordrecht: ICG Printing.

Hulk, A. (1986). Subject clitics and the pro-drop parameter. In Coopmans, Bordelois and Dotson-Smith (eds.), *Formal parameters of generative grammar II*. Dordrecht: Foris.

Ingham, Richard (1998). Tense without agreement in early clause structure. *Language Acquisition* 7(1):51–81.

Labelle, M., and D. Valois (1996). The status of postverbal subjects in French child language. *Probus* 8:53–80.

Lambrecht, Knud (1981). *Topic, antitopic and verb agreement in nonstandard French*. Amsterdam: John Benjamins.

MacWhinney, B. and C. Snow (1985) The Child Language Data Exchange System, *Journal of Child Language* 12, 271–96.

Nagy, Naomi, and William Reynolds (1994). Optimality theory and variable word-final deletion in Fætar. *Language Variation and Change* 9(1):37–55.

Prince, Alan, and Paul Smolensky (1993). *Optimality theory: Constraint interaction in generative grammar*. Ms., Rutgers University and University of Colorado Boulder.

Reynolds, William (1994). *Variation and phonological theory*. Ph.D. dissertation, University of Pennsylvania.

Rondal, J. A. (1985). *Adult-child interaction and the process of language acquisition*. New York: Praeger.

Rondal, J.A., J. F. Bachelet, and F. Peree (1985). Anayse du langage et des interactions verbales adulte-enfant. *Bulletin d'Audiophonologie* 5.

Suñer, M. (1988). The role of agreement in clitic-doubled constructions. *Natural Language and Linguistic Theory* 6:391–434.

Suppes, P., R. Smith, and M. Leveillé (1973). The French syntax of a child's noun phrases. *Archives de Psychologie* 42:207–269.

Schütze, Carson, and Wexler, Kenneth (1996). Subject case licensing and English root infinitives. *Proceedings of the Boston University Conference on Language Development 20*. Somerville, MA: Cascadilla Press.

Vainikka, Anne, Géraldine Legendre, and Marina Todorova (1999). PLU stages: An independent measure of early syntactic development. *Department of Cognitive Science Technical Report*, Johns Hopkins University.

Wexler, Kenneth (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106:23–79.