

Some questions collected from those handed in...

Concerning homework #7 and the OT syntax readings

To what extent is this a unified grammar if we're just making up constraints? And even if this part seems unified, doesn't it lead overall to a more fragmented grammar?

*Of course, no serious researcher is really trying to "just make up constraints." The way this is supposed to work is that one observes the patterns in the data, tries to make generalizations, and tries to get a feel for what the competing pressures seem to be that lead to the results we see. The first step might be just to give the pressures a name, but one also looks around in the larger environment for other people who have observed pressures that seem to be similar in an abstract way, and ideally one tries to unify the different pressures in terms of constraints. The more general the constraint, the more insight we've gained into how language works.*

*As to leading to a more fragmented grammar, this could be viewed in a couple of different ways. We don't know ahead of time what things in grammar are supposed to be grouped together and which ones aren't, so simply because looking at something in a new way changes the grouping doesn't necessarily mean it's not a step forward. The goal is always going to be to have the most general constraints possible—and in fact that's one of the selling points of OT. Because it is possible for constraints to be violated sometimes, they can be stated at their most general: Don't move, for example, or Telegraph. Of course, sometimes there **is** movement, and sometimes functional projections **are** pronounced. If we didn't allow for these constraints to be violated, then the constraints would have to be stated in a much less elegant way ("Don't move unless there is a strong uninterpretable feature that needs to be checked"). So, by looking at things in terms of OT, we can in certain situations get a better feel for what's going on.*

How is it possible to find the universal constraints without overgenerating or undergenerating the number of possible languages?

*That's precisely the difficulty—and the goal. If your constraints over- or under-generate, then this means that there's something you're missing. Either a language that would serve to exemplify the case your analysis predicts, or something about the constraints you are using. If the OT hypothesis is right, there **is** an answer to this question, and when we get there, we'll understand how language works and how languages can differ from one another. If, on the other hand, it isn't possible, then this suggests that the problem was in the OT hypothesis. We can't know without trying—and in trying, we also gain new insights into language by looking at the phenomena in new ways, even if OT in the end turns out not to work.*

What is the difference between having parameters and having rankings? Is there anything in OT like a “small, finite number” that we have with principles and parameters? It seems like there are an awful *lot* of constraints.

*Interesting observation. Yes, the direct analog in OT to “parameters” (like the head-parameter that differentiates SVO languages from SOV languages, or the null subject parameter that differentiates English and Italian) are rankings. So, languages differ (only) in terms of the different rankings they have between their constraints, and there are as many different kinds of languages as there are different rankings. So, it would appear that OT predicts **many** more language-types should be possible than principles & parameters approaches do. Of course, as we saw in the homework, it’s not the case that any two different rankings necessarily lead to two different languages (that can be distinguished empirically). Rankings between constraints that don’t conflict are generally not rankings that can be detected. Also, there are probably some rankings that are universally fixed: it’s always better for your syllables to have **a** as their nucleus than to have **k**, no language has that reversed. This too will cut down on the number of possible rankings a kid might be considering while acquiring a language. In the end, of course, the truth is what it is, our aesthetic sense of whether we like parameters better or rankings better doesn’t really have any bearing on it. This is where the “typology” predictions that OT makes are also important—all possible rankings of the constraints should be possible, so this is why when proposing an analysis, one should consider all of them to see if they’re all attested.*

What is the relation between PARMOVE, Superiority, the Minimal Link Condition, order preservation, and Attract Closest?

*These are all ways to express something that seems like a generalization about how language works. PARMOVE and order preservation here are in a somewhat different category—they express the generalization that language seems to have a bias against inverting the order of arguments. Superiority, the MLC, and AC, are all expressions of the generalization that language seems to minimize “effort” in some sense—movements are as short as they can be. How to spell out exactly what counts as “short” and what is and isn’t taken into consideration, that’s where they differ, and where the detailed research goes.*

If a resumptive pronoun is not inserted, that leads to an ungrammatical sentence (where the resumptive pronoun is there to repair an island violation). But then, how is that different from a Clash & Crash approach?

*I don’t think this has to be viewed as a Clash & Crash kind of phenomenon. It’s true that the resumptive pronoun is obligatory, meaning that a sentence that needs it would be ungrammatical without it. However, it relies on the comparison set—if you assume that you have a choice when starting the derivation between having the pronoun or not (for example, if you can either include it in your numeration at the beginning or not), then that’s a Clash & Crash case. If you needed it, and didn’t include it, the result is ungrammatical. However, if you assume that you **don’t** get to make that choice—that is, if your starting point is the*

*same whether the pronoun is pronounced or not, and that it's the grammar that decides whether a pronoun is pronounced there or not—then this is perfectly compatible with an OT approach where the best pronunciation wins. I don't think this phenomenon is one that really points away from OT—in fact, it has kind of an OT character, since it's a bit hard to state the conditions where a pronoun would be required in the numeration (or maybe more to the point, how a pronoun can save a structure) in a more Minimalist model.*

So, an object in Icelandic shifts when it has the feature F, but some NPs have that feature and some don't—so, in what way is this a reasonable explanation?

*It's basically a way to handle the apparent optionality in a framework that doesn't allow for optionality. It may not be elegant—and perhaps in the end, optionality might need to be allowed. But, on the other hand, we can now also be on the lookout for what F is. An analysis that posits a feature F that drives movement isn't claiming that the feature F has no grounding in reality. It could be, for example, that upon closer inspection, those things that move are interpreted as focused—in that case the F feature is something like [Focus].*

What is the Wackernagel position?

From Agbayani, Brian and Chris Golston (to appear). Second-position is first-position: Wackernagel's Law and the role of clausal conjunction. *Diachronica*.

*Based on the comparative syntax of early Indo-European (IE) languages, Delbrück (1878) and Wackernagel (1892) proposed that Proto-IE had a set of clitics which follow the first stressed word of their sentence. Their thesis has been so influential that second-position in a sentence is now usually referred to as Wackernagel's position, even outside of IE. Within IE, their conjecture has attained the status of a law. Watkins has gone so far as to say that "One of the few generally accepted syntactic statements about Indo-European is Wackernagel's Law, that enclitics originally occupied the second position in the sentence" (1964: 1036). Much contemporary work has sought to reevaluate the status of Wackernagel's Law for early IE languages within current linguistic theory. This body of recent work offers a descriptive precision that was not available to the frameworks within which Delbrück and Wackernagel worked; and this recent work has given rise to several analytic trends with respect to second-position (2ndP) phenomena in these languages [...]*

Wackernagel, Jacob. 1892. "Über ein Gesetz der indo-germanischen Wortstellung". *Indogermanische Forschungen* 1. 333-436.

Watkins, Calvert. 1964. "Preliminaries to the reconstruction of Indo-European sentence structure". *Proceedings of the 9th International Congress of Linguists* ed. by H. G. Lunt, 1035-1045. The Hague: Mouton.

Delbrück, Berthold. [ & Ernst Windisch]. 1878. *Syntaktische Forschungen*. Vol. III. Halle/S/ Max Niemeyer.

Another possible way that a sentence might crash could be that it just doesn't make any sense semantically, right? Would it be useful to look at an OT approach to semantics?

*True—it is quite possible that a senseless sentence could be syntactically well-formed, and semantic ill-formedness does provide an alternative way for unacceptable sentences to be ruled out grammatically. OT can also be applied directly to semantics, in ways similar to which it has been applied to syntax. OT*

*syntax is relatively new and not very deeply explored, and OT semantics is even newer and less deeply explored, but there are people considering the issue.*

In a language with no wh-fronting, is WHCRIT ranked very low or does it just not exist?

*It's a basic premise of OT that all of the constraints are in every language, and just the ranking differs. Partly, this is to explain how the constraints get there, because it seems rather impossible to actually **learn** what the constraints **are** in the process of acquisition (not to mention the poverty of the stimulus arguments about how the evidence isn't there, kids get to the target grammar reliably and quickly, ...).*

Why hasn't OT caught on in syntax? It's so nifty. Among those with whom it has caught on, is there any kind of consensus view on the role it plays in grammar?

*OT seemed to come to phonology at the right moment—people found that the phenomena they were looking at were more sensible, the explanations felt “deeper” with OT as a grammatical framework. The fact that it has swept the field so dramatically indicates that OT has opened a lot of avenues to understanding and explaining many phenomena in phonology that hadn't been available before thinking about things explicitly as constraint interactions. The situation seems to have been different with syntax. There is probably a sociological component to the lack of widespread use of OT in syntax, but there is probably also a lack of a comparable body of data that seems clearly more sensible from an OT perspective. Syntax doesn't seem to need OT in the same way phonology needed OT. In many cases, we understand things relatively well, and what OT brings is just an opportunity to “recode” what we already understood into constraints—without bringing a great deal more to the table. This is why Pesetsky, for example, went looking for places that did seem to have an “optimality character.” As for what use OT is put to, we've seen a couple and talked about a couple more, and I don't think that any one way is really dominant—there are small sets of researchers who work in each framework. Probably the most active area in which OT is used for syntax is actually in a different generative framework called LFG (Lexical Functional Grammar), for which Stanford is a primary home (but is used relatively widely outside of Stanford as well). LFG is rather different in many respects from the Minimalist syntax we've been working with; if you wish to look at it, see in particular work by Joan Bresnan.*

It seems hard to believe that in order to pronounce a sentence we need to construct a bunch of different possible trees in order to find the best one—isn't this too time-consuming?

*I think this is not the kind of assessment we can really make—first of all, is it really computationally expensive to make these comparisons? I think it depends on the way computation actually works—a step-by-step algorithm might not be ultimately the correct model. Perhaps everything is done in parallel. Second, what criteria can we apply to determine what would be too computationally expensive to be reasonable?*