

Syntax without syntax

Luke Smith

January 9, 2018

Abstract

Here I argue that the concept of narrowly syntactic parameters is unnecessary, and unbefitting of a Minimalist model of the language faculty. I attempt to describe an area of language classically thought of as being syntax-qua-syntax, that is, word order, and argue that the word order differences found in different languages can be said to be derived from differing prosodic constraints. To implement this, I craft an Optimality Theoretic account of the canonical word order of sentential constituents (the subject, object and verb), which closely approximates the real-world typology of existing languages, all motivated by phonological principles already existing, or with close analogs in the literature. That is, pre-existing prosodic constraints are sufficient to determine a language’s word order. I also show the enormous theoretical gains of this type of approach, noting the economy not just gained in theoretical simplicity, but in the clear account of how language is acquired by infants, that is, by a kind of robust phonological bootstrapping.

1 Background

Generative Grammar is often distinguished from its precursors or contemporary rivals as being a theory of innate structures. That is, while to pre-generative accounts of language, language is merely a complex behavior which is learned in a manner little different from other behaviors, after the Chomskyan revolution, it became a unique module with conceived as having idiosyncratic traits non-derivable from other cognitive modules. These traits, often summarized as “Universal Grammar” constituted the formal target of study for the Generative enterprise.

Indeed even a semantic change occurred around the word “grammar” itself. A “grammar” became a merely formal object in earliest accounts (Chomsky 1953, 1957, 1965), compared by Chomsky to a Post System. A grammar was a type of formal system, not necessarily a narrowly linguistic concept, although the core tenet of the Chomskyan program was that the distinguishing factor of the human language faculty was precisely the presence of a “grammar” which generated novel expressions from raw lexical material. On the more anthropological side, this generative ability was thought to have drastic effects of

human cognitive life, it is precisely this generative ability to which Berwick and Chomsky (2015) attribute human behavioral modernity and is what impels Chomsky (1966) to suggest that “linguistic and mental processes are virtually identical, language providing the primary means for free expression of thought and feeling, as well as for the functioning of the creative imagination.”

Still, the Generative tradition of grammar became overburdened with theoretical apparatus (arbitrary constraints, filters and stipulations) as the field advanced, prompting in the 1990’s a decidedly economy-oriented rejoinder to the field in the form of the *Minimalist Programme* (Chomsky 1995). The vantage point of the Minimalist Program would be that the complexity was not based on arbitrary constraints, but *emergent* based on interactions between “nature” and “nurture” and specifically through “laws of form,” the so-called “Third Factor” (Chomsky 2005).

Just conceptually, this solved the raw problem of theoretical economy, gave the field biolinguistic legroom: if the complexity of language emerges from laws of form, not complex programming, then we are not contradicting the known constraint of the peculiar rapidity of the evolution of the language faculty.

However, this creates a tension between this theoretical economy and the previous push for an autonomous formal grammar in the way described by Chomsky (1965). Some have taken this more traditional interpretation, arguing that the complexity of language is indeed deep and specific to the language faculty (Cinque 2012), ammunition against those who have spoken against the idea of a language faculty (Evans and Levinson 2009).

However a rigorously Minimalist interpretation of grammar requires us to variously discard what used to be thought of as being a component of “Universal Grammar,” replacing those constraints and stipulations with externally-motivated factors that drive those syntactic alternations traditionally thought of as being results of the narrow formal system of syntax.

Here I will argue for an extreme variety of Minimalism, in which some of the basic attributes traditionally ascribed to the narrow syntax, specifically, word order parameters, can be modeled entirely as an emergent reaction of interface constraints. While this can be construed as minimizing the reach of “Universal Grammar”, meaning that less of the domain of language is derived from language faculty-internal properties, this is a strong reinforcement of Minimalist principles.

So in Section 2 I will take this Minimalist intuition to the domain of word order, specifically, the ordering of the main constituents of a sentence, the subject, object and verb. I’ll argue that linguistic differences in word order are a function of prosodic differences, and we needn’t make reference to any formal syntactic layer of the language faculty to model these differences. I will specifically use an Optimality Theoretic framework to account for this in Section 3, and then will talk about many of the theoretical gains of this approach compared to the traditional pre-Minimalist account of word order in Section 5.

2 The Issue of Word Order

2.1 The Universality of Stress assignment

Word order parameters were one of the traditional examples of *syntax-qua-syntax*, a good example of the syntax of a language totally independent of any other fact. Still, this certainly shouldn't preclude any attempt to derive basic word order from extra-syntactic properties. Indeed, notionally, much of the talk in Minimalism is about deriving linguistic alternations from conceptually necessary distinctions at the levels of phonological or logical form.

At that, importantly, the traditional understanding of word order as an independent parameter doesn't adequately give us an accurate typology of actually-existing languages. The canonical word orders and the prosodic rules of human languages are arranged in a very particular way; all languages tend to have a prosodic structure in which places phrasal and intonational stress on the arguments of the verb, particularly an object or focal object in a clause (Gundel 1988).

Early generative attempts at typologizing phrasal stress, particularly the seminal Halle and Vergnaud (1987), treated phrasal stress as being "assigned" to words in a syntactic constituent based on parameters of a language, building gradually upwards to larger constituents. This was in keeping with the traditional Y-model of the grammar (Chomsky 1965), in which the syntax fed the phonological system a finished syntactic object, then presumably with a word order based on syntactic constraints.

Kahnemuyipour (2005) notices, however, that accounts like this over-generate the set of observable human languages. While Halle and Vergnaud (1987)'s theory is consistent with any word order with any stress pattern, Kahnemuyipour notes that observed human languages only take advantage of a subset of combinations of word orders and stress patterns, specifically, those that serve to produce languages in keeping with the generalization above: sentential stress should fall on the object and all arguments should be stressed over the verb.

Thus as a typological theory, Halle and Vergnaud (*ibid.*)'s account fails to sufficiently constrain the possible number of grammars, the surfeit of which yields a modeled language faculty with presumably more complexity than the actually-existing language faculty. Indeed, with some introspection, it should be clear that *any* theory in which syntactic constituents are ordered before and without reference to the prosodic tendencies will necessarily generate an unwanted over-abundance of possible prosodic grammars.

Now an ideal theory of word order would be one in which not only we can constraint the number of possible grammars to approximately what we observe in actually instantiated languages, but also a theory which motivates those parametric differences as being grounded in some kind of conceptually necessary external factor.

An obvious candidate for this motivating factor is the prosodic constraints themselves. That is, in language acquisition, we know that a child has to learn and prioritize the different prosodic constraints in language independent of what

else he learns. As mentioned before, it also seems that certain syntactic elements follow prosodic phrases in perfect lockstep. So instead of the classical model of the syntax deriving an order and consequently feeding that output into the phonological system, we can instead say that we first take the established phonological traits of a language and then use them to derive on an optimal word order given those constraints.

That way, our theory should avoid drastically over-producing phonologically ill-formed grammars in the way classical accounts had to, and it also greatly moves us toward a much more Minimalist understanding of the syntax of language, in that even the basic elements of a grammar are derived from conceptually necessary external constraints.

Our technical implementation for doing this will be Optimality Theory (Prince and Smolensky 1993), whose specifics will be explained in the ensuing section. It should be said that while Optimality Theory is still partially controversial, especially for use in syntax, I employ it here as a mere shorthand for the presumably much more complex computation performed in actually-existing neural nets, for more details, see Prince and Smolensky (1997). That is to say, my employment of Optimality Theory needn't entail a belief in a literal GEN or a mental tableaux, but only the use of the formalism as an indirect shadow imitating parallel computation in a neural net.

Indeed, this can be viewed as an extension of the general goal of Smolensky and Legendre (2006) to explicate linguistic and other cognitive phenomena as attempts to achieve an optimization of constraints or "Harmony". My goal here is specifically Minimalist, but precisely the tools given by Smolensky and Legendre (ibid.) allow us to posit an economical language faculty, leaving the heavy lifting and variation at the interfaces to more general cognitive architecture.

It should also be noted that while we have similar goals and methods as prosodically-attuned theories of syntax such as Contiguity Theory (Richards 2016), we differ in important ways. Firstly, Richards (2010, 2016) do indeed assume an incremental bottom-up construction of sentences by MERGE, but with an additional prosodic constraint requiring that probes and goals should be prosodically unified. Our assumption, rather, is that the language faculty needn't construct a sentence itself, but is in the business of *translation*, translating a representation in semantic space to an linearly-ordered prosodic space.

Additionally, unlike Richards (2016), we are not relying on traditional tree structure and derivations. That is, if we have an OSV sentence, I make no claim, or assent to the need for any claim about "what projection" the object is "in".

The only priors to this theory are bare essentials: subjects, objects, verbs (all semantically unordered) and prosodic constraints like stress and prosodic phrases with discrete borders. We make no claims about MERGE or any projections.

3 An Optimality-Theoretic Analysis

Again in order to move toward a plausible phonologically-grounded analysis, we can implement prosodic constraints in Optimality Theory (Prince and Smolensky 1993) and see what kind of word order typology can come about. In the following analysis, we will provide a constraint system two main forms: a transitive sentence (containing a subject (S), object (O) and verb (V)) and an intransitive one (containing only an S and V). Thus the input to the phonological system will not be a linearized string, similar to Halle and Vergnaud (1987), but an unordered set of constituents (in notation, indicated as being in {}).

The Optimality Theoretic system will then generate candidates for each one, each with either a different linear order or a different pattern of phonological phrasing uniting the constituents. After that, we will reorder and apply different constraints (listed and explained in Section 3.3) and analyze which of the theoretically-possible word order arise from this framework.

As a reminder, the main theoretical gain here is the lack of need for reference for independent syntactic parameters and a causal theory of word order that does not wildly over-generate grammars. That is, a theory like this where over-arching prosodic constraints drive basic word order, we not only give a causal account of existing data, but account the *absence* or impossibility of non-observed language patterns.

3.1 Generation of Candidates

We will say that GEN generates candidates that vary two main ways. First, the linear order of constituents; each logical order of the elements is viable. This means that as a starting point, we have the six logical possibilities of combinations of the subject, object and verb, i.e.: SOV, SVO, VSO, VOS, OVS, OSV.

Secondly, GEN can generate candidates with various phonological constituency. That is, each of the elements (S, O and V) must be put in (and may share) a phonological phrase. This means for each of the six orders, we have four different non-vacuous candidates, say for the SVO order, we can parse the constituents as [SVO], [S][VO], [SV][O] or [S][V][O].¹

3.2 Additional Assumptions

Left out is the possibility of a constituent to be “extra-metrical.” That is, GEN does not generate candidates such as S[VO] or [SO]V where one element is left unprosodified. For now, this is an issue of convenience, although it might make for an expanded typology for a further analysis.

Additionally it will be assumed in line with the aforementioned typological generalization that subjects and objects will always receive a stress superior to

¹There is the logical possibility of generating candidates such as [V][S][O][], etc. with empty phonological phrases. I will not include candidates like this, but they are dealt with by the to-be-mentioned * Φ constraint.

[SOV]	[S][OV]	[SO][V]	[S][O][V]
[SVO]	[S][VO]	[SV][O]	[S][V][O]
[VSO]	[V][SO]	[VS][O]	[V][S][O]
[VOS]	[V][OS]	[VO][S]	[V][O][S]
[OVS]	[O][VS]	[OV][S]	[O][V][S]
[OSV]	[O][SV]	[OS][V]	[O][S][V]

Figure 1: All theoretically possible word orders and parses for transitive verbs

[SV]	[S][V]
[VS]	[V][S]

Figure 2: All four possible parses for intransitive clauses

that of the verb, objects receiving the most stress. Therefore, implicitly, all constituents *already* have stress on them, thus [S][VO] is merely a shorthand for [Š][VÓ] (where the object takes the highest stress, followed by the subject, followed by the verb). GEN, therefore, does not generate candidates of the stripe [S][V̇O], etc. that violate the stress universals behind sentence constituent stress.

This may seem like begging the question, but I’d argue that in reality it is precisely how a theory of prosodic word order should be formulated. To make it clear, one can look at it this way: we can replace O, S and V with H, M and L respectively, for high, mid-level and low stress. That is, instead we can think of the OT derivation as deciding what points in the sentence sentential stress will be placed and phonological phrase boundaries will be formed.

And then “after” candidate-selection, the language faculty will seamlessly map the object onto the highest stressed position, the subject on the mid-level and the verb on the low or stressless position. The placement of these elements at the same levels of stress across all languages is precisely the constant behind word order differences. What varies is where that stressed position ends up in the sentence, and it is that which our constraints will determine.

All in all, we can sum up all of the possible candidates generated by GEN given these assumptions. Figure 1 shows all possible candidates for a transitive clause, there being 24—four possible parses for each of the six possible linear configurations of the subject, object and verb. Figure 2 shows the paltry four possible candidates for an intransitive clause.

3.3 Justification and Explanation of Constraints

Now that we’ve established the possible candidates, we can detail the constraints that will decide amongst them for particular grammars. We will be dealing with 8 total constraints, each one in one way or another derived from previous phonological literature or plausible typological generalizations.


$\{S, O, V\}$	CONT	IAMB
 a. [S][OV]		
b. [S][VO]		*!
c. [SO][V]	*!	
d. [SOV]	*!	
e. [SVO]	*!	*

Figure 3: CONT and IAMB exemplified.

3.3.1 Cont

The CONT (contour) principle requires that there be only one stressed element in each phonological phrase. This is a general prohibition against stress-clash, but needn't apply only to situations when the two stresses in question are directly adjacent.

In the implementation, a candidate violates CONT if both the stressed S and O appear within the same phonological phrase. Thus [SO][V] and [SVO] violate the constraint, while [S][OV] and [S][VO] do not.

3.3.2 Troc and Iamb

TROC and IAMB are similar and countervailing phonological constraints, now applied to the phrasal level. TROC demands that a constituent have trochaic stress (starting at the first subconstituent and alternating). And IAMB requires that a constituent have iambic stress (starting at the second subconstituent and alternating). Similar constraints have been variously posited in the literature, such as Selkirk (2011)'s STRONGSTART or Fitzgerald (1994)'s TROCHEE, both of which demand particular kinds of syntactic constituents to appear in crucial phrasal positions.

In this implementation, TROC and IAMB should be understood as applying to the whole intonational phrase (here effectively a sentence), and *not* the phonological phrases delineated with brackets: TROC and IAMB.

For example, all verb-initial languages VSO, VOS, in all possible parses, will violate TROC. All verb second languages (SVO, OVS) will violate IAMB.

3.3.3 Init Φ and Fin Φ

INIT Φ and FIN Φ are twin (mostly) countervailing constraints that desire a stressed element at either side of a phonological phrase. This kind of constraints could be rephrased if the actual phonetics requires it. That is, stress may be an epiphenomenon of more abstract constraints, such as where language faculty can construct a phrasal boundary *à la* Richards (2010).

In practice, INIT Φ incurs a violation whenever the unstressed V appears at the left edge of a phonological phrase. FIN Φ is violated when V appears at the end of a phonological phrase.


$\{S, O, V\}$	INIT Φ	FIN Φ
 a. [OV]		*
b. [VO]	*!	

Figure 4: A higher ranking of INIT Φ over FIN Φ favors clauses with initial stress.

Thus, [S][VO] violates INIT Φ , [OV][S] violates FIN Φ , [SVO] violates neither, and [S][V][O] both.

It’s important to sever the dichotomy between TROC and IAMB from INIT Φ and FIN Φ . While INIT Φ and FIN Φ are sensitive only to phrase edges, TROC and IAMB can be violated by any improperly metrical subconstituent. Additionally, FIN Φ , unlike IAMB is computed from the right edge.

3.3.4 Top1st

TOP1ST is the incarnation of the general functional tendency of languages to put old or aforementioned information as early (temporally) in a sentence as possible, with newer, focal information coming afterward.

It should be additionally noted that this constraint has strong typological corollaries, specifically that the overwhelming majority of languages prefer for topical elements (subject) to precede focal information (objects). Dryer (2013)’s tally shows that 1148 out of 1188 languages considered to have a canonical word order show a subject-before-object order (nearly 97% of languages).

In this implementation TOP1ST will yield a violation for any candidate that places its object to the left of its subject, thus all OSV, OVS and VOS orders sustain a violation of this constraint, while all SVO, SOV and VSO orders do not.

One may think of this in discourse or pragmatics terms, rather than simply prosody. If we abstract from the actual constituents and look only at the stress that associates with them, this constraint is violated whenever a primary stress (represented by O) occurs before a secondary stress (represented by S).²

3.3.5 * Φ

* Φ is simply a principle of economy applied to phonological phrasing. *Ceteris paribus*, a language will want to economize on the number of phonological phrases employed in any given structure.

In the implementation, a candidate incurs one * Φ violation for each phonological phrase it has. [OSV] will have one violation; [O][SV] will have two; [O][S][V] will have three, while a totally unparsed OSV would hypothetically have zero.

²This isn’t to say I have any problems modeling the computation of syntax, prosody, discourse relevance and everything else in language simultaneously. Indeed, I think such a synchrony is necessary.


$\{S, O, V\}$	CONST
a. [SO][V]	*!
 b. [S][VO]	

Figure 5: CONST wants two elements in a φ only when they are a semantic constituent.

It should be noted that $*_{\Phi}$ is solely responsible for weeding out vacuous candidates that I have not included in this analysis. It is $*_{\Phi}$ that rules out, say [V][SO][] as an alternative to [V][SO].

3.3.6 Const

The CONST principle represents the desire of the language faculty to map semantic or logical structures onto phonological structure. In simple terms, CONST will shun any form that incorporates elements into a phonological phrase which are *not* a constituent together.

In the context of subjects, verbs and objects, this means that two of these constituents are placed in the same phonological phrase, they must be the object and the verb, otherwise CONST yields a violation.

So [S][OV] incurs no violation, as the second phonological phrase containing the object and verb are a logical/syntactic constituent. [SO][V] however, does incur a violation, as the subject and object (which are not a logical constituent) are placed in the same phonological phrase without the verb. [VS][O] is similarly aberrant, as the subject and verb do not form a constituent.

To be clear, [VSO] or any other order of elements in one and only one phonological phrase does *not* incur a violation, as the whole phrase is indeed a constituent.

The fact that I have made reference to logical or syntactic constituency may seem like a kind of cheat given the fact that I said these constituents (S, V, O) could be thought of as merely different levels of stress. If that is the case, however, it should be clear that this constraint can be reformulated in phonological terms. That is, if two phonological constituents are to be under the same phonological phrase, they should be maximally distinct in terms of stress level. This makes this constraint, in phonological terms, almost an equivalent of CONT at a different level of abstraction.

Now that we've outlined the constraints, we can use these in an Optimality Theoretic analysis and discover what kind of word orders fall out from merely these constraints.

4 Implementation and Analysis

As mentioned in Section 3.1, the candidate set for transitive clauses consists of 24 viable candidates, that is, for each of the six logical subject, object, verb

Constraint	Description
CONT	* for φ containing both S and O
TROC	* for V initial parse
IAMB	* for second-position V parse
INIT Φ	* if V initial in φ
FIN Φ	* if V final in φ
TOP1 ST	* if O > S linearly
* Φ	Violation for every φ
CONST	* for every φ with non-logical constituent

Figure 6: Summary of constraints

orders, four different parsings, e.g. for the order SVO: [SVO], [S][VO], [SV][O] and [S][V][O] (see Figure 1).

The candidate set for intransitive clauses is significantly smaller, consisting only of the only four logically-possible parsings of the subject and verb alone: [SV], [S][V], [VS] and [V][S] (see Figure 2).

Given these possible word orders and parses, and given these constraints, the core question thus arises: **What possible word orders can be produced by these constraints given what candidates violate them?** Ideally, if we have constructed a decent theory, our constraints should be able to account for much of typological reality in actually-existing languages, while ruling out or dispreferring poorly-attesting orders.

It should also be noted that by “word orders” we mean pairs of transitive and intransitive orders. We might expect a language (i.e. ordering of constraints) that favors SVO clauses to also prefer SV intransitive clauses, also that is not a logical necessity, as languages such as Spanish demonstrate.

4.1 Methods

For ease of analysis, and in aid of creating a typology, I fed a spreadsheet of all candidates, constraints and the violations each constraint would incur for each candidate into the OT-Help program (Staubs et al. 2010).³ This software package takes such spreadsheets and returns an interactive report of possible languages resulting from different constraints interacting over different underlying forms.

This software package can be fed tableaux of constraints, candidates and their violations and will determine what possible constraint orders are consistent to produce the given candidates. Specifically, if there are word orders or parsings (from our list of 24 possible parsings for transitive clauses and 4 possible for intransitives) which are never possible for any given order of constraints, OT-Help will make us aware of that. Additionally, for each ordering of constraints,

³The source file used for this analysis can be found at http://luksmith.xyz/ling/word_order.csv

[SOV]	[S][OV]	[SO][V]	[S][O][V]
[SVO]	[S][VO]	[SV][O]	[S][V][O]
[VSO]	[V][SO]	[VS][O]	[V][S][O]
[VOS]	[V][OS]	[VO][S]	[V][O][S]
[OVS]	[O][VS]	[OV][S]	[O][V][S]
[OSV]	[O][SV]	[OS][V]	[O][S][V]

Figure 7: The possible word orders given our constraints

there will be precisely one transitive parsing and one intransitive parsing which will be consistent⁴. These pairings are important for our typology as well, as it may be that a certain type of intransitive clause necessitates another type of transitive clause or *vice versa*.

4.2 The Results and Typology

Given the 24 different candidates for transitive clauses and the four candidates for intransitive clauses, we could have 96 distinct languages, however, only ten of those, according to the analysis of OT Help can be produced by an ordering of constraints consistent across both transitive and intransitive clauses.

These ten possible language types are listed below by word order:

1. **An SVO/SV language**, with two subtypes, one where all constituents are in the same phonological phrase: [SVO]/[SV] and another where the transitive verb phrase is a phonological phrase unto itself [S][VO]/[SV]. The latter seems to closely approximate the English situation.
[SVO]/[SV]: CONST, * Φ , INIT Φ , TOP1ST, TROC > CONT, FIN Φ , > IAMB
[S][VO]/[SV]: CONT, CONST, TOP1ST, TROC > * Φ , FIN Φ > IAMB, INIT Φ
2. **An SVO/VS language**, that is, an SVO-type language whose intransitive clauses are verb-initial. This can be related to languages like Spanish. Again, two subtypes: one parsed [SVO]/[VS] and another parsed [S][VO]/[VS].
 - [SVO]/[VS]: CONST, * Φ , FIN Φ , TOP1ST > CONT, INIT Φ , TROC > IAMB
 - [S][VO]/[VS]: CONT, CONST, FIN Φ , TOP1ST > IAMB, * Φ , INIT Φ , TROC
3. **An SOV/SV language** with two subtypes, analogous to the SVO/SV language: one with all constituents under one phonological phrase [SOV]/[SV] (Basque-like), and another with the transitive VP as a phonological phrase alone: [S][OV]/[SV] (Persian-like).

⁴That is, barring circumstances where we do not have sufficient violations to distinguish candidates.

- [SOV]/[SV]: IAMB, CONST, $*_{\Phi}$, INIT Φ , TOP1ST, TROC > CONT, FIN Φ
 - [S][OV]/[SV]: CONT, IAMB, CONST, INIT Φ , TOP1ST, TROC > $*_{\Phi}$, FIN Φ
4. **An SOV/VS language** where the transitive VP is alone: [S][OV]/[VS]. This type is still mysterious to me, as I know of no language which seems to fit it well. There is only this one parse available, however.
- [S][OV]/[VS]: CONT, IAMB, CONST, TOP1ST > $*_{\Phi}$, FIN Φ > INIT Φ , TROC
5. **A VSO/VS language** of two types: [VS][O]/[VS] and [VSO]/[VS].
- [VS][O]/[VS]: CONT, IAMB, FIN Φ , TOP1ST > CONST, $*_{\Phi}$, INIT Φ , TROC
 - [VSO]/[VS]: IAMB, CONST, $*_{\Phi}$, FIN Φ , TOP1ST > CONT, INIT Φ , TROC
6. Lastly, **a VOS/VS language** parsed as [VO][S]/[VS].
- [VO][S]/[VS]: CONT, IAMB, CONST, FIN Φ > $*_{\Phi}$, INIT Φ , TOP1ST, TROC

Thus of the six logically possible *linear* word orders (ignoring phonological phrasing⁵), our constraints have yielded only four of them, precisely those four which make up $\approx 99\%$ of actually existing languages (1173 out of 1188 surveyed according to Dryer (2013)). Of the most common word orders (SOV/SVO/VSO), there are multiple parsing types.

4.3 Exemplification of and Comments on Constraint System

To make the analysis clear, we can examine some actual tableaux produced by this constraint system. We can first look at some individual language examples to show how bad candidates are ruled out of the system. After that, it's worth it to discuss some of the tendencies of the typology generally, and why they occur with the given constraints. First, let's implement an English-like [S][VO] language which has a [SV] structure in intransitive sentences.

4.3.1 An English-like [S][VO] language

In the transitive case (Figure 8), this word order violates both IAMB (because the V appears where iambic stress would otherwise be) and INIT Φ because the phonological phrase [VO] begins with that same unstressed verb. Thus an English-like language ranks both of these quite lowly.

⁵While it would behoove us to have wide typological data on phonological phrasing, to my knowledge, it is much harder to come by, and usually divided in many local analyses of single languages

$\{S, O, V\}$	CONT	CONST	TOP1 ST	TROC	* Φ	FIN Φ	IAMB	INIT Φ
☞ a. [S][VO]					**		*	*
b. [S][OV]					**	*!		
c. [SVO]	*!				*		*	
d. [SV][O]		*!			**	*	*	
e. [OV][S]			*!	*	**	*	*	
f. [SVO]	*!				*		*	
g. [VSO]	*!			*	*			*

Figure 8: The constraint ranking and tableau of an English-like SVO language.

$\{S, V\}$	CONT	CONST	TOP1 ST	TROC	* Φ	FIN Φ	IAMB	INIT Φ
☞ a. [SV]					*	*		
b. [S][V]					**!	*		*
c. [VS]				*!	*			*
d. [V][S]				*!	**	*		*

Figure 9: The constraint ranking and tableau of an English-like SVO language.

In the intransitive tableau (Figure 9), we see that [SV] violates the FIN Φ constraint, while its main competitor, [VS] does not, while violating TROC, thus meaning that TROC must be ranked more highly than FIN Φ . As a side note, it should be noticed that the constraints CONT, CONST and TOP1ST are all unviolated and inviolable by any hypothetical intransitive sentence.

4.3.2 On languages with only one phonological phrase

Minutely different from English-like [S][VO] languages are the [SVO] type of language. What differentiates these is simply a different ranking of CONT (which must be higher in English-like languages) and INIT Φ or * Φ , which would be higher in [SVO] languages.

This is usually generalizable to other word orders, what determines whether to parse a transitive sentence as either monophrasal or biphrasal is whether the language prioritizes phonological phrase economy (* Φ) or keeping stressed elements in different phonological phrases (CONT).

4.3.3 The difference between left-headed and right-headed structure

Given the English tableaux in Figures 8 and 9, we can compare this ranking with that of English’s right-headed [S][OV] analog, with the transitive and intransitive tableaux for the language in Figure 10 and 11 respectively.

Perhaps the core difference in “headedness” between [VO] and [OV] languages is the difference in ranking between the twin constraints INIT Φ and FIN Φ . The elements of the verb phrase are put together in such a way to satisfy

$\{S, O, V\}$	CONT	IAMB	CONST	INIT Φ	TOP1 ST	TROC	* Φ	FIN Φ
☞ a. [S][OV]							**	*
b. [SOV]	*!						*	*
c. [S][VO]		*!		*			**	
d. [VS][O]			*!	*		*	**	
e. [VO][S]				*!	*	*	**	

Figure 10: Transitive word order in an SOV language.

$\{S, V\}$	CONT	IAMB	CONST	INIT Φ	TOP1 ST	TROC	* Φ	FIN Φ
☞ a. [SV]							*	*
b. [S][V]				*!			**	*
c. [VS]				*!		*	*	
d. [V][S]				*!		*	**	*

Figure 11: Intransitive word order in an SOV language.

whichever of these constraints is more highly ranked. [OV] VPs are inherently stress initial, while [VO] VPs are stress final.

As a note, notice that, similar to in the case of SVO languages, what allows a [S][OV] parsing as opposed to a monoclausal [SOV] parsing is CONT being more highly ranked than * Φ . If these two constraints are rearranged, we get the constraint structure for a [SOV]/[SV] language.

4.3.4 The impossibility of one φ per constituent

For any of the six possible linear word orders, no valid language exists in this typology in which each of the three constituents are parsed in a phonological phrase alone, e.g., while [S][VO] and [SVO] are both possible parses, [S][V][O] is not, nor is any other ordering.

The reason for this is multifold. Firstly, these orders will always take the maximum three violations of * Φ , but aside from that, placing a V in a phonological phrase all alone means a violation of both INIT Φ and FIN Φ .

More generally, the framework here might be thought to conclusively rule out any verb-alone parsing, simply because any such parse could undergo *some* kind of Pareto improvement by additionally parsing *any* adjacent constituent. If we expanded our theory to make room for extrametricality, we might allow for parses such as [SO]V, where V is extrametrical, although I will not pursue that further here for the time being.

4.3.5 The (near) lack of O > S order

Of the ten theoretical parsings, only one allows for the object to appear linearly before the subject: [VO][S]/[VS]. As stated before from Dryer (2013)’s analysis, this is by far the most common of the still vanishingly rare O > S orders, and

in our system it seems to be a perfect storm of missed ranking violations as we can see below. Such an order arises from a ranking as follows:

CONT, IAMB, CONST, FIN Φ > * Φ , INIT Φ , TOP1ST, TROC

The highest four constraints happen to wipe out all of [VO][S]’s competitors, while the parsing violates all four of the lower ones. While this order is permitted in the grammar given, other O > S languages are variously ruled out. TOP1ST, by definition, is of course a weight against all of these orders.

A note on emergence of candidates. It might seem like TOP1ST is fundamental in ridding our typology of O > S orders. This is surprisingly *not the case*. If we run the same candidates with otherwise the same constraints⁶, we get a typology quite similar to the one in this analysis with only the one same O > S parse [VO][S], *but notably lacking VSO orders*. TOP1ST does not by itself serve to weed out unwanted orders, but allows to the existence of an optimal VSO solution. A way to think about this is that there are many possible constraint orderings which without TOP1ST would generate a [VO][S] grammar, which is a local optimum. While TOP1ST does not remove this local maximum, it redirects some possible grammars to a nearby local maximum, either [VS][O] or [VSO], depending on the particular constraint order.

4.4 Review of analysis

We now have a colorful and reasonably accurate typology of the world’s existing languages with specific predictions of what possible parsings may occur. I don’t at all pretend that this analysis is perfect or anything other than a feeble beginning, nor do I doubt that similar work with a slimming or streamlining of constraints or assumptions, but let’s make the general point clear.

The great boon of constraint-based analysis and the more general idea that the mind is a harmonic optimizer gives us this (1) enormous empirical coverage or individual and typological linguistic facts (2) given few, but plausible derived assumptions (3) without having to posit additional mental organs and UG-machinery.

This allows us to have a truly Minimalist account of the origin of not just language facts, but of typological facts as well. At that, classical account of syntax-prosody, such as (Halle and Vergnaud 1987), which put the cart (syntax) before the horse (prosody) are troubled by the constant need for somewhat arbitrary stipulation to avoid overgeneralizing the possible set of grammars.

5 Formal and functional extensions

In fact, kind of analysis of language is highly pliable to new empirical domains. Specifically, because it links syntactic word order to non-syntactic fact, not only can we make predictions about a language’s syntax by noting the traits of its

⁶For the source file for this analysis, see http://lukesmith.xyz/ling/word_order_wo_topf.csv.

phonology, but we can also establish causal, empirical links between otherwise unlinkable syntactic “parameters.”

Let’s take the generalizations made by work such as Greenberg (1963). Many of them relate corollaries across different syntactic categories. For example Universal 17 states that “with overwhelmingly more than chance frequency, languages with dominant order VSO have the adjective after the noun.” A fact like this might be viewed as “butterfly-collecting” from some formal vantage points, but if we model syntax as falling out from tangible prosodic parameters, it suddenly becomes an important relation. While there might not be anything formal or semantic binding these two facts, it might be easy to state that these sentential and NP word orders both arise when particular *phonological* parameters are particularly highly or lowly ranked.

Or perhaps even morphological universals; take Universal 27: “If a language is exclusively suffixing, it is postpositional; if it is exclusively prefixing, it is prepositional.” The symmetry between morphology and adpositions in language *could* be simply stipulated in the syntax, and perhaps reasons as being part-and-parcel in a distributed approach to morphology, but we can with no more difficulty say that this fact falls out from prosodic constraints which favor unstressed elements be on the right or left side of a heavier constituent. This makes a more falsifiable statement about the phonological corollaries of the situation, but also provides a non-stipulated reason for why there should be a uniformity between two theoretically distinct (although prosodically similar) elements. Certain languages may tend care greatly about particular constraints, which may affect parts of the phonology, parts of morphology and parts of syntax.

This allows us not only to make predictions of individual languages, but also of how languages will tend to fall into linguistic types, or why it is often observed that languages often diachronically are shooting for a typologically harmonious symmetry of headedness parameters or other macro features.

In fact, if we failed to interweave prosody, morphology and syntax in positions like this where they seem to be causally connected, traditional more linear models of grammar will fail to account for why the syntactic engine succeeds in constructing precisely that structure which is needed by the phonological interface, the same core problem we ran into with general word order and stress assignment.

But a rejiggering of theory in a prosodically-driven direction not only increases the possibility of scientific falsifiability (by making tangible claims of what prosodic traits a particular language should have), but it also substantially widens the domain of linguistics beyond what is common in the Generative Program: allowing us to make both typological and diachronic predictions.

5.1 Headedness parameters

Present in many of Greenberg’s universals and near-universals, as well as present in common linguistic parlance is the concept of the two general language types “head-initial” and “head-final.” While languages can vary as to what kinds of

syntactic categories take what kinds of orders, languages often cluster around abiding my general types like these. For example, VSO languages are overwhelmingly likely to place adjectives before the noun and seem to all (according again to Greenberg (1963)) use prepositions rather than postpositions and place auxiliaries before main verbs.

In isolation, this statement is a kind of formal coincidence in traditional syntax: Why should all these languages happen to have adpositional, auxiliary and headedness parameters that all harmonize? But once we widen the view include external factors, such as prosodic rules, as potential causal factors, we can see some symmetry. It can easily be said that in the same way a language orders the subject, object and verb in such a way to stress the lexically and discursively important arguments over the more grammatical verb, it might do the same in the ordering of adpositions or adjectives.

To simplify things, a “head-initial” language might simply be one with a generally low ranking of TROC or INIT Φ , and a correspondingly high ranking of FIN Φ . This will generally locate stress on the right side of phrases, thus making it optimal across all syntactic categories to place more prominent and lexical elements to the right of their heads. This thus brings us a much more plausible formal apparatus for accounting not just for individual languages, but for macro-linguistic tendencies.

5.2 Language acquisition as a multi-dimensional puzzle

Linking prosodic and syntactic information in such a way also opens a new dimension by which infants acquire language. In traditional conceptions, the infant is faced with the dual problems of learning the phonology of a target language on one hand, and the morphosyntax on the other. There are a set of syntactic parameters a child must learn, and there is a non-overlapping set of phonological rules and parameters to be learned.

This means that a child must battle on several fronts simultaneously, and the different features of phonology, syntax, and all other aspects of language are additive in difficulty.

But in our new conception however, not only are prosody and syntax “the same,” but they are mutually reinforcing. That is, prosodic data, even without any lexical knowledge, hold phrasal divisions with stress in particular places in such a way that it can be used for phonological bootstrapping to make more informed assumptions about the lexical categories that presumably appear there. A child can hear a sentence with a particular intonation and assume not only where the stress goes, but what kind of word (in lexical category) should be there.

Some experimental work, in the domain of sentential word order, seem to validate this idea; Grünloh, Lieven, and Tomasello (2011), for example, find that German children rely most on *prosody*, not even context, when determining whether a sentence is SVO or OVS, the most important clue to the role of the first element is not its meaning, but the *pitch on the word itself*.

It’s clear also that the relevance of prosody to sentence comprehension is indeed immediate (see studies such as Eckstein and Friederici (2006), Kerkhofs et al. (2007), and Sammler et al. (2010), all of which show immediate interaction of prosody on how a sentence is interpreted). A corollary of this is that prosodic cues are one of the main factors in “disambiguating” sentences; specifically, if unaware of the presence of ambiguous sentences, speakers produce utterances with the intonation proper to exhaustively communicate syntactic category, even if the other listener is equally unaware of the ambiguity (Millotte, Wales, and Christophe (2007) find precisely this in experimental circumstances).⁷ I put “disambiguate” in scare quotes as this “ambiguity” only exists once you write a sentence down—otherwise prosodic detail is sufficient to preclude ambiguity.

Prosodic processing is concomitant with “syntactic” processing in these cases, and the acquisition of both advances simultaneously in first language acquisition (Männel and Friederici 2011).

Again, what this means is that while language acquisition may be a puzzle with many dimensions, syntactic and prosodic, every new data point in prosody sheds light on syntax, and *vice versa*. Instead of acquiring hundreds of different language parameters in different domains, infants are merely narrowing in on a constraint order illustrated by every data point of both syntax and phonology.

6 Differences from other broadly “phonological” accounts

Our account is not dissimilar to Hammond (2011)’s, where it is argued that prosodic structure is generated by a grammar, and words and morphemes are chosen among alternatives based on adherence to that prosodic structure. The impetus for this is similar to ours here: while there’s no sense in which the syntax conditions phonology, there are relatively clear cases of the reverse, where a prosodic environment accounts for the selection of a particular allomorph or lexical item. Therefore Hammond (ibid.) accounts for the same problem we noted in Halle and Vergnaud (1987): when word order is a function of syntactic parameters, we will overgenerate the number of possible grammars, while if we assume that the phonology provides a set of well-formed prosodic templates in which words are inserted, we do not.

Schlüter (2015) introduces an overtly evolutionary/diachronic model off of the same intuition, arguing that when grammars have multiple synonymous forms (i.e. *sunk* vs. *sunken*) they gradually move to selecting one or the other in particular prosodic environments depending on which will avoid marked metrical structures. This is distinct from our analysis here in that we are arguing that the selection of forms, particularly word order is part of the synchronic “grammar”, or more clearly a third-factor property of a given ranking of constraints.

⁷Similar results are found by Bögels et al. (2009) in the context of control sentences as well. Even in situations of apparent ambiguity, prosodic detail gives speakers all that is necessary to understand sentences that would be ambiguous on paper.

7 Closing

All of this we should expect if, as I have argued, syntactic parameters *are* merely results of prosodic parameters. Languages vary in what interface constraints they prioritize (their constraint rankings) and these constraints are emergent from external limitations of the externalization scheme of speech. This kind of account of syntactic difference is thus a movement to a maximally Minimal language faculty, where variation, absent at the level of the language faculty itself, is present only at the interfaces.

To repeat, given these several constraints, we can closely approximate the word orders of $\approx 99\%$ of the world’s languages, and we can plausibly motivate these constraints in such a way that makes tangible, falsifiable predictions about what sort of prosodic characteristics should correlate with what word orders.

It also opens to gates to similar analysis branching over syntactic category, phrasal level and even macro-comparison, an area of linguistics which has been somewhat out of the reach of the formal tools of Generative Linguistics thus far.

References

- Berwick, Robert and Noam Chomsky (2015). *Why Only Us: Language and Evolution*. MIT Press.
- Bögels, Sara et al. (2009). “The Interplay between Prosody and Syntax in Sentence Processing: The Case of Subject- and Object-control Verbs”. In: *Journal of Cognitive Neuroscience* 22.5, pp. 1036–1053.
- Chomsky, Noam (1953). “Systems of Syntactic Analysis”. In: *The Journal of Symbolic Logic* 18.3, pp. 242–256.
- (1957). *Syntactic Structures*. Mouton & Co.
- (1965). *Aspects of the Theory of Syntax*. MIT Press.
- (1966). *Cartesian Linguistics*. Harper & Row.
- (1995). *The Minimalist Program*. MIT Press.
- (2005). “Three Factors in Language Design”. In: *Linguistic Inquiry* 36, pp. 1–11.
- Cinque, Guglielmo (2012). “Cognition, universal grammar, and typological generalization”. In: *Lingua* 130, pp. 50–65.
- Dryer, Matthew S. (2013). “Order of Subject, Object and Verb”. In: *The World Atlas of Language Structures Online*. Ed. by Matthew S. Dryer and Martin Haspelmath. Max Planck Institute for Evolutionary Anthropology.
- Eckstein, Korinna and Angela D. Friederici (2006). “It’s Early: Event-related Potential Evidence for Initial Interaction of Syntax and Prosody in Speech Comprehension”. In: *Journal of Cognitive Neuroscience* 18.10, pp. 1696–1711.
- Evans, Nicholas and Stephen C. Levinson (2009). “The myth of language universals: Language diversity and its importance for cognitive science”. In: *Behavioral and Brain Sciences*.

- Fitzgerald, Colleen (1994). “Prosody drives the syntax: O’odham rhythm”. In: *Proceedings of BLS 20*.
- Greenberg, Joseph H. (1963). “Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements”. In: *Universals of Language*. Ed. by Joseph H. Greenberg. MIT Press.
- Grünloh, Thomas, Elena Lieven, and Michael Tomasello (2011). “German children use prosody to identify participant roles in transitive sentences”. In: *Cognitive Linguistics* 22.2, pp. 393–419.
- Gundel, Jeanette K. (1988). “Universals of topic-comment structure”. In: *Studies in Syntactic Typology*. Ed. by Michael Hammond, Edith A. Moravcsik, and Jessica R. Wirth. John Benjamins Publishing Company.
- Halle, Morris and Jean-Roger Vergnaud (1987). *An Essay on Stress*. MIT Press.
- Hammond, Michael (2011). “Phonology is the generative component”.
- Kahnemuyipour, Arsalan (2005). *The Syntax of Sentential Stress*. Oxford University Press.
- Kerkhofs, Roel et al. (2007). “Discourse, Syntax, and Prosody: The Brain Reveals an Immediate Interaction”. In: *Journal of Cognitive Neuroscience* 19.9, pp. 1421–1434.
- Männel, Claudia and Angela D. Friederici (2011). “Intonational phrase structure processing at different stages of syntax acquisition: ERP studies in 2-, 3-, and 6-year-old children”. In: *Developmental Science* 14.4, pp. 786–798.
- Millotte, Séverine, Roger Wales, and Anne Christophe (2007). “Phrasal prosody disambiguates syntax”. In: *Language and Cognitive Processes* 22.6, pp. 898–909.
- Prince, Alan and Paul Smolensky (1993). *Optimality Theory: Constraint interaction in generative grammar*. University of Boulder.
- (1997). “Optimality: From Neural Networks to Universal Grammar”. In: *Science* 275, pp. 1604–1610.
- Richards, Norvin (2010). *Uttering Trees*. MIT Press.
- (2016). *Contiguity Theory*. MIT Press.
- Sammler, Daniela et al. (2010). “Prosody meets syntax: the role of the corpus callosum”. In: *Brain* 133, pp. 2643–2655.
- Schlüter, Julia (2015). “Rhythmic influence on grammar: scope and limitations”. In: *Rhythm in Cognition and Grammar*. Ed. by Ralf Vogel and Ruben van de Vijver. Mouton de Gruyter.
- Selkirk, Elisabeth (2011). “The Syntax-Phonology Interface”. In: *The Handbook of Phonological Theory*. Ed. by John Goldsmith, Jason Riggle, and Alan C. L. Yu. Blackwell.
- Smolensky, Paul and Géraldine Legendre (2006). *The Harmonic Mind: From Neural Computation to Optimality-theoretic Grammar*. MIT Press.
- Staub, Robert et al. (2010). *OT-Help 2.0*. University of Massachusetts Amherst.