The Structure of Syntactic Typologies

JANE GRIMSHAW

Abstract: This article illustrates how language variation and the limits of variation are given a shared and principled explanation in Optimality Theory. It shows that languages can be ‘uniform’, choosing the same grammatical structures in three different sentence types. They can also be ‘non-uniform’, but the combinations of grammatical structures that they can exhibit are extremely restricted. The theory characterizes possible and impossible grammatical systems without special stipulations or additional theoretical machinery.

1. Introduction

At the heart of all research in generative grammar is the thesis that grammars are cognitive systems of the human mind, shaped by both genetic endowment and experience. The research program of generative grammar is committed to the view that grammars are universal in a profound sense, being constructed from universally available elements and components which are interconnected, and interact, according to principles of linguistic theory. The principal challenge facing linguistic theory is to explain simultaneously the massive variation among languages and their overwhelming similarity. The grammars of individual languages are neither arbitrary, nor entirely universal in any obvious, immediately observable sense. Language variation must be explained in a way that is compatible with the fact that children, equipped with universal grammar plus whatever other resources they recruit in successful language acquisition, are able to learn all of the observed linguistic variants, or ‘languages’. Developing a theory of universal grammar, explaining language variation and explaining how language acquisition is possible, are three facets of the same project.

The project is remarkably challenging: just how challenging becomes apparent when we study any particular linguistic phenomenon from this perspective. The fact that theoretical linguists engage in such analysis on a daily basis while others,
including most researchers in psychology and philosophy do not, goes a long way toward explaining the widespread lack of understanding of the theory of universal grammar outside theoretical linguistics. The evidence for hypotheses of the generality intrinsic to a theory of universal grammar stems from deep understanding of many highly specific instances of linguistic phenomena. The theory of what is universal about grammar necessarily goes hand in hand with a theory of what can be particular to individual grammars. It is just as important to understand and explain how grammars vary as it is to understand and explain how they do not.

Consequently, within the research program, two clusters of questions are the foci whenever any empirical linguistic phenomenon is investigated. First, what aspects of the phenomenon are attributable to universal principles, and what are those universal principles? Second, what aspects of the phenomenon do not reduce to universal principles, but rather reflect alternative grammatical systems at work, beneath the umbrella of (and thus consistent with) universal grammar? What principled properties are reflected in the patterns of variation among grammars, and what theory explains these?

From this perspective, a successful theory of universal grammar must necessarily also be, or provide for, a successful theory of variation—a theory of what is not universal, but is governed by universal grammatical principles. A theory which makes no predictions about language variation does not meet the challenge of explaining human mastery of language. Such theories can at best treat language variation as an inconvenience to be accommodated.

The present article offers a demonstration of how language variation and the limits of language variation are given a shared and principled explanation in Optimality Theory (Prince and Smolensky, 2004). The argument is as follows. (i) The same structural options are available in all grammatical systems. (ii) Grammars choose which of the universally available options they employ. (iii) A grammar can choose different options in different grammatical circumstances. I refer to this as ‘non-uniformity’. (iv) The theory predicts that some, but not all, combinations of options are possible. The choice of one option in one circumstance can determine or limit the choices available in another. To put it another way, the typology, i.e. the set of grammatically possible languages, is highly structured, and variation is greatly restricted.

In Optimality Theory (‘OT’) the principles that construct the set of alternative (competing) configurations, and the set of constraints which evaluate those configurations are both universal, i.e. they are exactly the same in every grammar. Under this theory, the analysis of a particular phenomenon in an individual grammar has immediate consequences for all other grammars and the languages that they generate. If a constraint is required in the analysis of any language it is in the grammar of all languages. If an option is grammatical in any language then it must be among the set of competing alternatives (‘candidates’) available in all grammars.

The constraints are universal and are hypothesized to be innate. Language variation is due to alternative ranking of the universal constraints. Ranking must
be learned.¹ (See Tesar and Smolensky, 1998; Tesar to appear; and much research available on the Rutgers Optimality Archive for background and recent results.) The ranking of the constraints is what determines which of the universally available options a grammar picks as grammatical in a given circumstance.

Variation in ranking predicts the variation that we should find in natural languages, other things (such as learnability) being equal. Here I explore one facet of variation: the existence and properties of systems which exhibit non-uniformity, as sketched above. The argument is that all and only those non-uniform systems which result from possible rankings of the set of universal grammatical constraints are predicted to be possible. Non-uniform systems, in fact all systems, which require impossible (i.e. inconsistent) rankings of the universal constraints are perforce themselves impossible. The theory thus entails that not every logically possible way of combining structures results in a possible human linguistic system, imposing principled limitations on language variation.

The two other theories of grammar represented in this collection do not provide any intrinsic insight into how languages combine structures. Neither Construction Grammar (see Goldberg, 2006, 2009), nor theories based on parameter setting or the choice of feature values (such as work stemming from the ‘Principles and Parameters’ model of Chomsky, 1981), or the Minimalist Program initiated in Chomsky, 1995, address how the choice of a parameter setting, the value of a syntactic feature, or the nature of a particular construction in a language relates to other choices in the grammar of the same language. It does not follow from these theories that there should be any relationship at all between the setting of one parameter and the setting of another, or the value of a feature in one configuration and the properties found in another configuration. Statements bearing on the relationship between e.g. the grammar of Negation and the grammar of VP-displacement can be added to a theory of course, but principled relationships between them do not follow from the fundamental principles of the theories themselves. This point is developed in the concluding section.

The design of this article is intended to make it possible to read and understand the general argument, which is presented in Section 2, without grappling with the formal grammar which underlies it. The formal analysis is presented in Section 3, along with a small set of examples. Grimshaw, 2013, presents a more rigorous and detailed presentation of the proposals summarized here.

2. Typologies and Their Structure

Under analysis here are three types of tensed clause, each having three alternative versions. Each of these alternatives (or ‘candidates’) is grammatical in some language.

¹ I set aside here the status of morphemes, which to some extent must also be learned.
The typological question is which combinations of grammatical candidates are possible in a single language. This defines the set of language types.

Consider a situation in which there are three sets, and each contains three members:

(1) \{A, B, C\} \{I, II, III\} \{! @ #\}

Each language consists of one member from each set, in the simple case. Now the question arises: is every logically possible combination a possible language? The answer which research in generative grammar, including the present work, strongly supports is ‘No’. Some combinations are possible and some are impossible. The question is what combinations of choices are possible and why.

In OT the same theoretical mechanism, namely constraint ranking, defines what is a possible winning candidate, and what is a possible combination of winning candidates, that is, the typology of possible languages. The key theoretical tool which controls both of these results is the set of universal constraints and the theory of optimization.

A possible system is one in which there is a consistent ranking of the universal constraints which selects the desired member from each of the sets, yielding a set which defines a possible language \{B, III, !\}, for example. An impossible system is one that cannot be defined by a consistent constraint ranking: selection of the desired member from one set requires that some constraint dominate another constraint, while selection of the desired member from another set requires the opposite domination relationship. So a grammar which chooses B may be unable to choose II.

Five universal constraints, with their effects governed by the OT theory of evaluation, are responsible for the typological predictions investigated here. When ranked, they form the grammar (in the relevant respects) of the individual languages. At the same time, they determine what variation within and among languages is possible and what is not. This is the logic of OT and of the present argument.

Informally, one of the universal constraints requires that any phrase have a head; one requires that tense and a verb must be realized in adjacent positions; one bans meaningless elements; one disallows main verbs in the layer of structure above the normal main verb position; one disallows main verbs in the complementiser (‘C’) position, which is in a yet higher level of syntactic structure. The constraints are defined more precisely below in (19).

The three configurations which form the paradigm to be analyzed are illustrated by the English examples in (2)–(4).

(2) She \textbf{did}n’t read the book. \hspace{1cm} \textbf{[Negation]} \\
(3) Which book \textbf{did} she read? \hspace{1cm} \textbf{[Wh-question]} \\
(4) They said that she \textbf{would} read the book, and read the book she \textbf{did}. \hspace{1cm} \textbf{[VP-displacement]}

The configurations which languages can choose among are themselves defined by universal grammar, and consist of the structures which respect universal well-formedness principles. For example, in most OT work it is hypothesized that they must conform to ‘X-bar theory’, the core theory of syntactic structure. (See Carnie, 2012, for a current textbook treatment.)
A few points are of particular interest here. Some form of the verb *do* appears (bolded) in all of the structures in (2)–(4). (This phenomenon is usually referred to as ‘do-support’ and is here abbreviated as ‘DS’.) The English paradigm is *uniform*, in the sense that DS occurs in all three of the configurations. Another point concerns the relationship between the constraints and the paradigm. The constraints are highly general, and make no reference to the configurations in (2)–(4). Nevertheless, as we will see, they make very strong predictions about which choices can combine within a single grammatical system, and which cannot. Languages like English, which choose DS in all three configurations, are predicted to be possible, but certain other combinations are not. Hence the constraints entail strong restrictions on the types of languages that can exist.

Why do we find DS in English in these configurations? Part of the answer is that Negation, Wh-questions and VP-displacement pose a particular set of challenges to the system of Universal Grammar. They separate the verb from Tense (‘T’), which is ultimately realized as part of the verb in many languages, which is possible only if T is, approximately, adjacent to the verb in the syntactic representation. An analysis of DS along these general lines has been in circulation since Chomsky, 1957. In (5) T is separated from the verb, due to the presence of an intervening *not*.

(5) Negation

```
      TP
     /   \
  she   T'

     /   \
  T   VP

      /  \
 Past not VP

      /  \
  V   NP

    /  \
read the book
```

In (6) the separation is due to the fact that the VP is displaced leftwards.

(6) VP-displacement

```
      CP
     /   \
 VP   C'

     /   \
read the book C

     /   \
  C   TP

    /  \
  she T'

     /  \
  T   VP

      /  \
 Past V NP

    /  \
Ø   Ø
```
In (7), the empty C position must be filled, so Tense must move into it. Here again, Tense is separated from V.

(7) Wh-question

\[
\begin{array}{c}
\text{CP} \\
\text{which book} \\
\text{C'} \\
\text{C} \\
\text{TP} \\
\text{she} \\
\text{T'} \\
\text{T} \\
\text{Past} \\
\text{V} \\
\text{NP} \\
\text{read} \\
\emptyset
\end{array}
\]

The fact that the structures in (5)–(7) separate V from T provides part of the explanation for why we find DS in (2)–(4). The second part of the explanation is that DS is the solution to the challenge which is chosen by the grammar of English.

While Negation, Wh-questions and VP-displacement are the focus here, it should be emphasized that many other grammatical situations can engage DS, and that separation of T from the verb is just one of the occasions for DS within English and cross-linguistically. DS is also used to ‘verbalize’ clausal heads which are not verbal, to realize ‘verum focus’ and for many other purposes. Far from being specific to particular constructions, it provides a solution for a wide variety of grammatical challenges. DS is also observed in various circumstances in a wide variety of languages from many language families for which we have evidence. See the references in Grimshaw, 2013, for a hint of the larger picture.

Many languages, which face the same challenges as English because they are similar enough in the relevant respects, do not avail themselves of DS in counterparts to (2)–(4). Instead they utilize other grammatical options, two of which are focused on here. One alternative places a main verb in the position which do occupies in each of the configurations in the paradigm in (2)–(4). In accord with the extensive literature stemming from Emonds, 1978, and Pollock, 1989, in this option a main verb raises to a higher position in the clausal structure. I will refer to the resulting structures as involving ‘Verb Raising’, abbreviated as ‘VR’. Both the DS solution and the VR solution arrange the clause so that Tense is adjacent to a V after all.

Yet other systems leave the problematic T/V configurations unaltered, using neither do nor a main verb there, settling for a tense morpheme with no adjacent verb. I will refer to this structure as an instance of ‘Free Tense’, abbreviated as ‘FT’. The languages I will examine here do not include any FT examples, but Grimshaw, 2013, offers such an analysis of will, drawing on research presented in Elías-Ulloa, 2001, a work which proposes tense-stranding in the Panoan language Shipibo.

Returning to the set-theoretic characterization in (1), we can now elaborate it to display the typological choices under discussion. Negative clauses, Wh-questions
and clauses involving VP-displacement each allow for three different versions: VR, DS and FT.

(8) Negation: \{DS, VR, FT\}
    Wh-questions: \{DS, VR, FT\}
    VP-displacement: \{DS, VR, FT\}

Each set of alternatives is a ‘candidate set’ in OT. Each member of the candidate sets in (8) is the option chosen by some language, i.e. the optimum under some ranking of the universal constraints. Now we can ask which combinations of the members of the candidate sets form possible grammatical systems, and which do not.

Since each candidate set has three members, the logically possible grammatical systems number 27: 3x3x3. However, under the principles of Optimality Theory, the selection of members from the three candidate sets is not independent. Certain choices are predicted to be incompatible with other choices. The interaction of the constraints thus entails that, among the 27, only certain combinations constitute possible human linguistic systems; the remainder do not. I emphasize again that none of the constraints refers to Negation, Wh-questions or VP-displacement, let alone to combinations of them within a single language. The typological predictions result from interaction among the constraints, not from the definition of the constraints themselves. These typological predictions are laid out in the remainder of this section and explored with greater precision in the following section.

At the two extremes are systems that I will call ‘(maximally) uniform’ and ‘(maximally) non-uniform’. Uniform systems make the same choices for each of the three candidate sets. There are three uniform systems, given in (9). Six systems are maximally non-uniform: they pick a different option in each of the three candidate sets. These are given in (10).

English, as we have seen informally, is a uniform language, choosing DS in all three columns in (9) in the past and present tenses. The future tense as analyzed in Grimshaw, 2013, is an example of a uniform system with FT in all cases, exemplifying the third row of (9).

(9) Negation Wh-question VP-displacement Language

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VR</td>
<td>VR</td>
<td>VR</td>
<td>?</td>
</tr>
<tr>
<td>DS</td>
<td>DS</td>
<td>DS</td>
<td>English past/present</td>
</tr>
<tr>
<td>FT</td>
<td>FT</td>
<td>FT</td>
<td>English future</td>
</tr>
</tbody>
</table>

The constraints posited here, interacting as OT requires, predict that these 3 uniform systems are possible systems. While it might seem that predicting uniform systems is

---

2 The first row is a complex case to evaluate because the violation of c-command which results from VR and VP-displacement in the same structure may prevent this structure from being possible at all. See Grimshaw, 2013, for some relevant work and references.

© 2013 John Wiley & Sons Ltd
trivial, that they will inevitably arise as a result of whatever analysis is provided for each component candidate set, this is false. As Section 3 will establish, uniformity results if and only if there is a consistent ranking which delivers a uniform choice of optima across all candidate sets under consideration. This is the case for the systems in (9), which are therefore predicted to be possible language types.

The maximally non-uniform systems are impossible: there is no consistent ranking of the five constraints at work which will pick a different member from each of the three candidate sets. (See (19) for the most explicit version of the constraints given here.) Hence all of the ‘languages’ (combinations of optima) in (10) are predicted to be impossible as human languages, although they are, of course, logically possible.

(10)

<table>
<thead>
<tr>
<th>Negation</th>
<th>Wh-question</th>
<th>VP-displacement</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>VR</td>
<td>DS</td>
<td>FT</td>
<td>–</td>
</tr>
<tr>
<td>VR</td>
<td>FT</td>
<td>DS</td>
<td>–</td>
</tr>
<tr>
<td>DS</td>
<td>VR</td>
<td>FT</td>
<td>–</td>
</tr>
<tr>
<td>DS</td>
<td>FT</td>
<td>VR</td>
<td>–</td>
</tr>
<tr>
<td>FT</td>
<td>DS</td>
<td>VR</td>
<td>–</td>
</tr>
<tr>
<td>FT</td>
<td>VR</td>
<td>DS</td>
<td>–</td>
</tr>
</tbody>
</table>

While maximally non-uniform combinations are impossible according to the theory, this does not mean that all non-uniform systems are impossible. Two prominent empirically established examples of non-uniformity within the paradigm are German and Monnese, which will be the principal instances of non-uniformity considered here. Monnese is a Northern Italian dialect analyzed in Benincà and Poletto, 2004, in which there is a discrepancy between the Negation candidate set and the Wh-question candidate set: VR is the choice for Negation, while DS is the choice for Wh-questions (see Benincà and Poletto, 2004, pp. 70–71). The verb *fa ‘do’ cannot appear in negated clauses, where VR is found instead, as illustrated by Benincà and Poletto’s (26) (pp. 70–71). The raised verb and the alternative *fa ‘do’ are both bolded for ease of recognition.

(11)  
(a) l so mia
      It I know not
(b) *fo mia savè-l
      ‘I do not know-it’

Yet in Wh-questions, *fa-support is grammatical and VR is not. (12) is Benincà and Poletto’s (1b, d) (p. 52).

(12)  
(a) ke fa-l majà?
      What does-he eat
(b) *ke majà-l?
      What eats-he
We can characterize Monnese informally by saying that *do*-support is preferred when it avoids a violation of the constraint against main verbs in C (*LinA*), even though it is not preferred when just a violation of the constraint against main verbs above their normal position (*LinF*) is at issue. Hence Monnese does not tolerate a lexical verb in C even though it tolerates one in T.

The existence of a system with VR for Negation and DS for Wh-questions provides us with the opportunity to test the predictions of the OT analysis. Indeed, despite predicting that maximally non-uniform systems are impossible, OT with the posited constraints predicts that (13) is a possible combination of structures.

(13)

<table>
<thead>
<tr>
<th>Negation</th>
<th>Wh-question</th>
<th>VP-displacement</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>VR</td>
<td>DS</td>
<td>—</td>
<td>Monnese</td>
</tr>
</tbody>
</table>

In German both Negation, as in (14), and Wh-questions as in (16), involve VR. (These examples were provided by Fabian Heck.)

(14) (a) Ich weiß es nicht
     I know it not
(b) *Ich tu(e) es nicht wissen
     I do it not know
     'I do not know it'

(15) (a) Was isst er?
     What eats he
(b) *Was tut er essen?
     What does he eat

Yet the optimum for VP-displacement is like English, and involves DS, as shown in Bader and Schmid, 2006. (15) is their (14a).

(15) Tanzen tut Katja immer noch häufig.
     Dance *does Katja still often

Hence the combinations for German are those in (16):

(16)

<table>
<thead>
<tr>
<th>Negation</th>
<th>Wh-question</th>
<th>VP-displacement</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>VR</td>
<td>VR</td>
<td>DS</td>
<td>German</td>
</tr>
</tbody>
</table>

3 Benincà and Poletto do not report VP-displacement in Monnese. They note (in Section 3.2.3) that VP ellipsis after auxiliaries is impossible. Since no evidence is available with respect to the VP-displacement column I leave it blank.
Why are non-uniform systems possible? Does a theory which predicts that they are possible necessarily admit all kinds of non-uniform systems, short of the maximally non-uniform cases already eliminated? The answer is that OT predicts the existence of some but not all non-uniform systems. Moreover, the division between the two is not random but reflects structure imposed on the candidate sets by the universal constraints, as we will see.

Like all systems, a non-uniform language is possible if and only if there is a consistent ranking of the proposed five constraints of UG which chooses different structures in different candidate sets. German and Monnese are both generated by the constraints interacting according to the principles of OT, which entail that both fall within the set of possible human linguistic systems. The constraints also entail that two hypothetical systems, which I call ‘Anti-Monnese’ and ‘Anti-German’, are not possible. Anti-Monnese uses DS with Negation and VR in Wh-questions, so it has DS where Monnese has VR, and VR where Monnese has DS. Similarly, Anti-German uses DS with Negation in Wh-questions and VR in VP-displacement. The Anti-Monnese and Anti-German choices are given in (17).

(17)

<table>
<thead>
<tr>
<th>Negation</th>
<th>Wh-question</th>
<th>VP-displacement</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS</td>
<td>VR</td>
<td>—</td>
<td>Anti-Monnese</td>
</tr>
<tr>
<td>DS</td>
<td>DS</td>
<td>VR</td>
<td>Anti-German</td>
</tr>
</tbody>
</table>

The reason why neither Anti-German nor Anti-Monnese is generated is that there is no ranking of the relevant universal constraints which derives the combination of choices in (17), whereas German, Monnese and English combine choices in a way which is compatible with at least one ranking of the universal constraints.

It is important to see that the constraints do not themselves encode any information about combinations of choices, or ‘optima’, at all. As mentioned earlier in this section the constraints do not encode the relationship between choices in one candidate set and choices in others. It is the interaction of the constraints, not the constraints themselves, which determines which combinations of structures are possible and which are not. In the following section I provide a more precise account of the constraints and discuss in more detail how they interact.

Each language can now be viewed as a set of best choices/optima: the set consisting of the chosen variant of Negation plus the chosen variant of Wh-question etc. The typology generated by any set of constraints interacting according to OT is the set of languages generated by them, i.e. the set of sets of best choices, or optima.

The five universal constraints and the candidate sets discussed here yield a typology which contains 9 languages, i.e. different combinations of grammatical structures for Negation, Wh-questions and VP-displacement. The typology is calculated using OTWorkplace (Prince and Tesar, 2012). These languages can be distilled into types, by collapsing into a single type those languages which use the same structure (VR, DS or FT) in a given candidate set, but which, for
reasons external to present considerations, place the V, or do or T in a different structural position. (See Grimshaw, 2011 and 2013, for the complete OTWorkplace calculations and the full (un-compressed) typology.)

The 6 non-uniform language types that are contained within the typology fall into three groups, represented in (18).

(18)

<table>
<thead>
<tr>
<th>Combination</th>
<th>Negation</th>
<th>Wh-question</th>
<th>VP-displacement</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>VR</td>
<td>VR</td>
<td>DS/FT</td>
<td>German</td>
</tr>
<tr>
<td>2</td>
<td>VR</td>
<td>DS</td>
<td>DS/VR</td>
<td>(Monnese)</td>
</tr>
<tr>
<td>3</td>
<td>VR</td>
<td>FT</td>
<td>FT/VR</td>
<td></td>
</tr>
</tbody>
</table>

This typology is highly structured. It allows for non-uniform grammatical choices, but it does not allow for all cases of non-uniformity. One group it allows has VR for Negation and Wh-questions, with divergence for VP-displacement. The other two groups choose VR for Negation, diverge for Wh-questions and then choose either the Negation structure or the Wh-question structure for VP-displacement.

Note that all of the non-uniform language types choose VR for Negation. The rankings required to choose the VR structure for Negation (given in (23) below) are compatible with the rankings required to choose VR, DS, or FT for Wh-questions, and for VP-displacement. However, the rankings required to choose either the DS structure or the FT structure for Negation force the same choice in Wh-questions and VP-displacement, and thus yield uniform systems. This is demonstrated for Monnese and Anti-Monnese Negation and Wh-questions in Section 3. It is more fully demonstrated in Grimshaw, 2011 and 2013.

The language types represented in (18), plus the uniform systems in (9), exhaust the combinations of structures generated by the theory, i.e. generated by the universal constraints under all rankings. Of the 27 logically possible combinations of forms, only 9 are predicted to be possible: 3 uniform and 6 non-uniform. While the number of possible rankings of 5 constraints is 5! (5x4x3x2x1 = 120), the number of distinct language types is 9.4

The two known non-uniform languages, German and Monnese, are generated by the system, and they appear in the typology as instances of combination #1 and #2 respectively. (The parentheses around Monnese in the Language column of (18) are intended to indicate that the VP-displacement cell is empty for Monnese, so it is not a perfect match to language type #2.) Anti-German and Anti-Monnese do not appear in (18). They both choose DS in the Negation column, but every

---

4 In the fullest system analyzed in Grimshaw, 2011 and 2013, there are 11 constraints and 6 candidate sets. The language types are broken into distinct languages by the location of the raised V, do, or free tense, and the tenses are considered separately. There are 26,244 logically possible combinations of optima, and the number of rankings of the constraints is almost 40 million. The system generates 162 languages.
language in the typology which chooses DS for Negation chooses DS everywhere, i.e. is uniform. Anti-German and Anti-Monnese are impossible languages, because they involve combinations of choices which cannot be generated by any ranking of the constraints.

In sum, universal constraints are ranked to form the grammar of individual languages. The ranking determines which constraints are satisfied in the best choice and which are violated in the best choice, where the ‘best choice’ or ‘optimum’ is the choice that is grammatical for the language.

Any combination of choices is possible if and only if there is a consistent ranking of the universal constraints which generates it. If there is no such ranking for a particular combination of optima, a language which exhibits the combination is impossible. The language is inconsistent with the theory of universal grammar as defined by Optimality Theory; it cannot be generated with the universal constraints posited here, interacting under the universal theory of optimization.

The following section contains a more technical presentation which shows how and why the proposed universal constraints, when ranked to form the grammars of individual languages, entail the typology just discussed. The reader who prefers to can proceed directly to the concluding section.

3. Constraints, Candidates and the Typology

In this section I provide more technical detail in support of the proposed explanation for how and why the non-uniform pattern of Monnese is generated and how and why the non-uniform pattern of Anti-Monnese is not. I will also demonstrate how and why the uniform system of English arises. The logic is essentially identical for all three patterns.

Recall that every grammar consists of the same set of constraints: the constraints are universal. ‘Markedness’ constraints assess output structures, and ‘faithfulness’ constraints assess the relationship between the input and the output structures. The constraints involved in the typology under discussion are those in (19). I have retained the constraint names used in the literature that the analysis is drawn from, but I have taken the liberty of modifying the definitions slightly, to make the analysis easier to follow.

(19) The constraints

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>V + T</td>
<td>Violated by a T which is not adjacent to a V (auxiliary or main)</td>
</tr>
<tr>
<td>ObHd</td>
<td>Violated by a phrase which has no head</td>
</tr>
<tr>
<td>FullInt</td>
<td>Violated by an element which has no meaning (here auxiliary do)</td>
</tr>
<tr>
<td>∗LinF</td>
<td>Violated by a main verb which occurs outside VP</td>
</tr>
<tr>
<td>∗LinA'</td>
<td>Violated by a main verb which occurs in C</td>
</tr>
</tbody>
</table>

5 These constraints, which are based on a wide variety of proposals by others, have their OT origin in Grimshaw, 1997.
The alternative candidates in each case are the three structures that might be utilized to realize inputs for Negation, Wh-questions and VP-displacement. The set of alternative realizations, the ‘candidates’, is universal. Every grammar must choose among the same set of candidates. DS, VR and FT are universally available, but they appear in grammatical sentences only in languages with rankings, i.e. grammars, which allow them in optimal candidates. The structures are represented abstractly in (20). This format shows whether the Tense is associated with a raised main verb (a.), with do (b.) or is free (c.).

(20) a. VR: [TP Subj V-T ... [VP <V> ...]]
b. DS: [TP Subj do-T ... [VP V ...]]
c. FT: [TP Subj T ... [VP V ...]]

The constraints are extremely general and are therefore widely violated. This is illustrated by (21), where the violations of the 5 constraints in the 3 alternative structures given in (20) are recorded. (21a–c) are instances of (20a–c) respectively. In English words, candidate a. corresponds to I do not know ... , b. to I do not know ... and c. to I past not know (the pattern found in the future in English: cf. I will not know.)

(21) Violation table: Negation

<table>
<thead>
<tr>
<th>Input: {V Subj T Neg}</th>
<th>V+T</th>
<th>OB HD</th>
<th>FULL INT</th>
<th>*LIN A</th>
<th>*LIN F</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. VR</td>
<td>[TP Subj V-T Neg [VP &lt;V&gt; ...]]</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. DS</td>
<td>[TP Subj do-T Neg [VP V ...]]</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. FT</td>
<td>[TP Subj T Neg [VP V ...]]</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The constraint violations that each candidate incurs are indicated in the column corresponding to each constraint. Each of the structures is preferred over the alternatives by at least one constraint and dis-preferred by at least one constraint. The preferences are the result of the patterns of constraint violation. The VR candidate violates *LinF, because the lexical head V has raised outside VP. Hence this constraint prefers the DS and FT candidates to the VR candidate. The DS candidate violates FullInt, since auxiliary do is meaningless, so this constraint prefers the VR and FT candidates to the VR candidate. The FT candidate violates V+T, which thus prefers both the VR and DS candidates to FT.7

6 Since the critical constraints are all markedness constraints, the details of the inputs are unimportant here, though not, of course, in general. I include schematic inputs, which simply indicate the principal grammatical components to be assembled in the candidates under discussion.

7 Some properties of violation tables throughout this article are due to the artificial limitations imposed by the nature of this report, and do not reflect general properties of constraint violation in OT: e.g. in (21) no candidate violates more than one constraint and no candidate
This small example illustrates many of the essential principles of OT. The candidates, the constraints and the constraint violations are the same in every language: the language merely chooses which of the available options it will employ. The constraints are maximally general and are massively violated. The grammatical candidate, the ‘winner’, is the one which best satisfies the constraints. It need not (and in practice never does) satisfy all of the constraints: it must be optimal, not perfect. The grammar of the language, by means of the ranking/prioritization of the constraints, chooses which candidate is optimal from each of the candidate sets defined by universal grammar.

To determine which candidate best satisfies the constraints, the candidates must be compared. The comparison is directly represented in a Comparative Tableau (CT) which exhibits the relationship between the winning candidate and the losing candidates, in a particular candidate set (Prince, 2002a, 2002b; Brasoveanu and Prince, 2011). This makes transparent which rankings are critical in choosing the winner over the losers. The CT for Monnese Negation is given in (22). The candidates are exactly those analyzed in (21). In (22), the winning VR structure is bolded and marked with ‘W’: see row a. of (22). Each row represents the preference of the constraints encoded by the columns, with respect to the winner candidate in the row.

There are three possibilities. A constraint may prefer the (bolded) winner, indicated by ‘W’ in the cell. It may prefer the loser, which is indicated by ‘L’. If the constraint does not distinguish between the winner and the other candidate, the cell is blank. ‘W’ in a cell thus indicates that the constraint heading the column prefers the winner to the candidate represented in the row. ‘L’ in a cell indicates that the constraint prefers the candidate itself (a losing candidate) to the winner.

(22) Negation VR winner

<table>
<thead>
<tr>
<th>Input: {V Subj T Neg}</th>
<th>V+T</th>
<th>OB</th>
<th>HD</th>
<th>FULL</th>
<th>INT</th>
<th>*LIN A</th>
<th>*LIN F</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. W VR</td>
<td>[TP Subj V-T Neg [VP &lt;V&gt; ... ]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. DS</td>
<td>[TP Subj do-T Neg [VP V ... ]]</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. FT</td>
<td>[TP Subj T Neg [VP V ... ]]</td>
<td>W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In (22) there is one constraint which prefers both of the losing candidates to the winner, namely *LinF, which is violated by the winner and satisfied by the two other candidates. There are two constraints which prefer the VR winner, violates a single constraint more than once. This is far from true of patterns of violation more generally, and is not true in the fuller analysis of Grimshaw, 2013.

8 Candidates may have identical violation combinations, so there may be more than one winner, but this is a situation that is not relevant for current purposes.
namely $V + T$ and \texttt{FULLInt}. Both of these constraints must rank above $^{*}\text{LinF}$ if the Monnese winner is to be optimal and hence grammatical. The notation ‘$\gg$’ is read as ‘dominate(s)’ or ‘rank(s) above’.

(23) **Monnese (Negation VR winner)**

\texttt{FULLInt} and $V + T \gg ^{*}\text{LinF}$

This is the first essential piece of the Monnese puzzle: the ranking in (23) picks the VR structure for Negation. But recall that in Monnese we find DS, not VR, in Wh-questions: Monnese is non-uniform. How does this come about? The violation table for Wh-questions is given in (24):

(24) **Violation table: Wh-questions**

\begin{center}
\begin{tabular}{|l|c|c|c|c|}
\hline
Input: \{V Subj T Wh\} & V$+$T & $^{*}\text{Lin}$ & $^{*}\text{Lin}$ & \texttt{FULL} \\
\hline
\texttt{HD} & \texttt{OB} & \texttt{INT} & A' & F \\
\hline
a. VR & \texttt{CP Wh V-T [TP Subj} & \texttt{<V-T>} [\texttt{VP} & \texttt{<V>} ...]] & 1 & 1 \\
b. DS & \texttt{CP Wh do-T [TP Subj} & \texttt{<do-T>} [\texttt{VP} ] & 1 & 1 \\
c. FT & \texttt{CP Wh T [TP Subj} & \texttt{<T>} [\texttt{VP} ] & 1 & 1 \\
d. $\emptyset$C & \texttt{CP Wh __ [TP Subj} & \texttt{V-T [VP} & \texttt{<V>} ...]] & 1 & 1 \\
\hline
\end{tabular}
\end{center}

Again using English words to represent the candidates schematically, candidate a. corresponds to *What ate he*, b. corresponds to *What did he eat*, c. to *What past he eat*, (the pattern found in the future in English: cf *What will he eat*), and d. to *What he ate*.

Here we see a candidate which we have not examined thus far, because no such candidate is relevant for Negation. The structure of Wh-questions like (12) must contain an extra phrase (usually referred to as ‘CP’) to house the fronted Wh expression, but \texttt{ObHd} will be violated unless this extra phrase has a head ‘C’.

(25) \[
\texttt{CP Wh [C C TP ]}
\]

The three familiar structures are available here: The C could be filled by a main verb, filled by \texttt{do}, or filled by Tense. But there is a fourth option, labeled as ‘OC’ in (24), which leaves the head position empty, thus violating \texttt{ObHd}. The violations incurred by this candidate are in (24d). There are thus four key candidates to be considered. The CT for the DS winner, the candidate which is grammatical in Monnese, is given in (26).

---

9 A candidate of this type in the Negation candidate set can never be the optimum because it is ‘harmonically bounded’ (Prince and Smolensky, 2004; Samek-Lodovici and Prince, 2005).

10 Here I only consider candidates in which \texttt{do} is in C, rather than in some lower head position such as T. These additional candidates are included in Grimshaw, 2011 and 2013.
(26) Wh-question DS winner

<table>
<thead>
<tr>
<th>Input: {V Subj T Wh}</th>
<th>V+T</th>
<th>ObHD</th>
<th>FullInt</th>
<th>*LinA′</th>
<th>*LinF</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. VR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP Wh V-T [TP Subj &lt;V-T&gt; [VP &lt;V&gt; ...]]]</td>
<td>L</td>
<td>W</td>
<td>W</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. W DS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP Wh do-T [TP Subj &lt;do-P&gt; VP ]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. FT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP Wh T [TP Subj &lt;T&gt; VP ]]</td>
<td>W</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. ØC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP Wh __ [TP Subj V-T [VP &lt;V&gt; ...]]]</td>
<td>W</td>
<td>L</td>
<td>W</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

By the same reasoning applied above, the CT for the DS optimum motivates the rankings in (27). FullInt prefers every loser over the winner, since the winner violates it while all of the losers satisfy it. The constraint ranking must therefore ensure that FullInt is not decisive. In order for the DS structure to be chosen over the VR structure (row a.), either *LinA′ or *LinF must out-rank FullInt. In order for the DS structure to be chosen over the FT structure in row c., V+T must out-rank FullInt. Finally, in order for the DS structure to be chosen over the structure with an empty head (row d.), either ObHd or *LinF must out-rank FullInt. We thus arrive at the rankings in (27):

(27) Monnese (Wh-question DS winner)

V+T ≫ FullInt
ObHd or *LinF ≫ FullInt
*Rina′ or *LinF ≫ FullInt

These rankings yield a non-uniform system, with VR in the case of Negation and DS in the case of Wh-questions. The VR winner for Negation was chosen by the ranking in (23). Comparison of the rankings in (23) and (27) shows that they are consistent, i.e. that there are total rankings of the constraints which are compatible with both. In fact the information provided by (23) resolves the disjunctions in (27): the Negation choice shows that FullInt must dominate *LinF, so we know that it must be ObHd and *LinA′ that dominate FullInt in Monnese, not *LinF. Any total ranking compatible with (28) will yield the non-uniform Monnese system.

(28) Monnese ranking derived from Negation and Wh-question candidates

V+T and ObHd and *LinA′ ≫ FullInt
FullInt and V+T ≫ *LinF

This is why the theory admits a non-uniform system with the properties exhibited by Monnese. There is a ranking of the universal constraints which generates both the DS choice in Wh-questions and the VR choice in Negation.

The same is not true for the system which I call ‘Anti-Monnese’, which has DS with Negation but VR with Wh-questions. No language can pick this combination of optima. The rankings which select the DS candidate for Negation, as in English (and Anti-Monnese), are those in (29), which are derived from the CT in (30), where the DS winner is in row b. Note that the candidates here are those in (21) and (22), only the choice of winner is different.
(29) Anti-Monnese/English (Negation DS winner)
\[ V + T \text{ and } ^*\text{LinF} \gg \text{FullInt} \]

(30) Negation DS winner

<table>
<thead>
<tr>
<th>Input: {V Subj T Neg}</th>
<th>V+T</th>
<th>OB HD</th>
<th>Full Int</th>
<th>^*Lin A’</th>
<th>^*Lin F</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. VR [TP Subj V-T Neg [VP &lt;V&gt; ... ]]</td>
<td></td>
<td></td>
<td>L</td>
<td>W</td>
<td></td>
</tr>
<tr>
<td>b. W DS [TP Subj do-T Neg [VP V ... ]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. FT [TP Subj T Neg [VP V ... ]]</td>
<td>W</td>
<td>L</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Choosing the VR candidate for Wh-questions requires that FullInt dominate \(^*\text{LinF}\) and \(^*\text{LinA’}\), as can be seen from (31), where the VR winner is in row a. and the DS loser in row b. (In order to simplify the presentation, the FT and OC candidate are omitted from (31); the remaining candidates, schematized using English words, correspond to What ate he and What did he eat.) The crucial observation is that in order for the VR candidate to win over the DS candidate, FullInt must dominate \(^*\text{LinF}\) and \(^*\text{LinA’}\), both of which prefer the DS candidate.

(31) Wh-question VR winner

<table>
<thead>
<tr>
<th>Input: {V Subj T Wh}</th>
<th>V+T</th>
<th>OB HD</th>
<th>Full Int</th>
<th>^*Lin A’</th>
<th>^*Lin F</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. W VR [CP Wh V-T [TP Subj &lt;V-T&gt; [TP &lt;V&gt; ... ]]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. DS [CP Wh do-T [TP Subj &lt;do-T&gt; VP ]]]</td>
<td>W</td>
<td>L</td>
<td>L</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(32) Anti-Monnese Wh-question VR winner: FullInt \(\gg^*\text{LinA’}\) and \(^*\text{LinF}\)

The grammar of Anti-Monnese must combine the rankings in (29), which pick the DS winner with Negation, with the ranking in (32), but these are obviously inconsistent. In (29) \(^*\text{LinF}\) dominates FullInt, but in (32), FullInt dominates \(^*\text{LinF}\). It follows that Anti-Monnese is not generated by this constraint system under the OT theory of constraint interaction.\(^{11}\)

In contrast with Anti-Monnese, English chooses DS with Negation, and also chooses DS with Wh-questions and VP-displacement: it is a uniform system with

\(^{11}\) It has been suggested to me informally on more than one occasion, that the regulation of movement of a lexical verb to C is managed by the ‘Head Movement Constraint’ (Travis, 1984), which prevents a head from moving across a filled head position. But the critical VR candidate, which loses to DS in Monnese but would win in Anti-Monnese, contains no intermediate filled head position to block movement (through T) to C, so the HMC cannot block movement of V to C in Monnese. Indeed Benincà and Poletto (2004) do not propose that the HMC is responsible.
The Structure of Syntactic Typologies

respect to the typology under consideration. A uniform system which requires an inconsistent ranking is impossible, for exactly the same reason as Monnese. Both uniform and non-uniform systems are possible only if they are generated by at least one total ranking of the constraints. To see why it is possible to choose the DS option for both Negation and Wh-questions, consider (33). (33) includes the same candidates as (31); but the winning choice is now DS rather than VR.

(33) Wh-question DS winner

<table>
<thead>
<tr>
<th>Input: {V Subj T Wh}</th>
<th>V+T</th>
<th>OB</th>
<th>FULL</th>
<th>*LIN A</th>
<th>*LIN F</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. VR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP Wh V-T [TP Subj &lt;V-T&gt; [vp &lt;V&gt; … ]]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. W DS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[CP Wh do-T [TP Subj &lt;do-T&gt; VP ]]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FullInt prefers the losing VR candidate, but both *LinF and *LinA’ prefer DS. So any ranking in which one of these two constraints dominates FullInt will choose candidate b. The ranking for a DS winner with Negation, in (29), already demands that *LinF ≫ FullInt, so this ranking is compatible with that required by DS in Wh-questions.

In fact, the two rankings are more than just compatible: the ranking in (29) entails that either *LinF or *LinA’ dominates FullInt. To put it another way, the ranking which chooses DS for Negation is compatible only with the choice of DS for Wh-questions, and also in fact, although this is not demonstrated here, for VP-displacement. The choice of a candidate as optimal in one candidate set has consequences for the selection of optima in other candidate sets. (See Anttila and Andrus, 2006; Prince, 2006, for research on entailment relations of this kind.) In contrast, the choice of VR with Negation is compatible with many alternative rankings, which can choose different optima in different candidate sets. This is the pattern that was observed in the typology presented in Section 2.

The English ranking enforces the same choice across all three candidate sets. The Monnese ranking enforces different choices for different cases. Thus the theory predicts the possibility of uniform choices, as well as the possibility of non-uniformity. But both are possible only if there is a consistent ranking of the universal constraints which picks the uniform or non-uniform combinations of optima.

4. Conclusion: The Limits of Variation Within and Among Languages

Under OT, variation among languages is inevitable: it is the result of alternative rankings of universal constraints. Non-variation is equally inevitable. It arises in circumstances where alternative rankings of constraints give the same result. Here I have argued that syntactic variation and non-variation, in the guise of what I have called ‘uniformity’ and ‘non-uniformity’ are inevitable within languages.
They arise in exactly the same way and for exactly the same reasons as (non)-uniformity across languages. Variation within languages, not a traditional focus of typological research, is governed under OT by the same principle as variation among languages, namely the ranking of universal constraints (Prince and Smolensky, 2004). Language-internal non-uniformity is both inevitable and highly constrained: some non-uniform grammars are possible and the rest are not.

From the perspective of learning, a non-uniform system is the same as a uniform system: the ranking of the universal constraints which will yield the system needs to be determined. Any candidate set where the choice of optimum critically involves a particular ranking serves as evidence: a learner can set the ranking based on the CT for that candidate set. Learning non-uniform systems is no more difficult, from this perspective, than learning uniform systems. Both involve determining the required rankings of the universal constraints. Learning non-uniform systems like Anti-Monnese or Anti-German is impossible, since there is no ranking of the universal constraints which will generate them.

In order to explain the patterns of variation analyzed here, both within and between languages, it is essential that the choice among alternatives, such as VR, DS and FT, are not made once for the language as a whole. This would yield only uniform systems. Neither are they made separately for each candidate set, or ‘construction’. (This predicts 27 distinct grammars as noted in Section 2.) If each candidate set or construction is associated with its own ranking, every logically possible combination of VR, DS and FT, including Anti-Monnese and Anti-German, would be generated.

OT makes it inevitable that grammars make choices among potential winning candidates (grammatical structures) neither at the level of the entire grammar, nor at the level of the individual candidate set. Since the constraint ranking which constitutes a grammar is fixed for all competitions, it governs the choice of the optimum in all candidate sets. Candidate sets cannot vary arbitrarily in their choices. But the fixed ranking of the constraints which governs any individual language does not necessarily impose uniformity across candidate sets, as we have seen. A ranking can, for example, pick VR for Negation and DS for Wh-questions. Still, non-uniformity is reined in by the effects of ranking; there is no ranking that can pick DS for Negation and VR for Wh-questions, as demonstrated here.

Under the principles of OT, cross-linguistic variation and invariance are explained in the same way as within-language uniformity and non-uniformity. They follow from the same grammatical theory. That both are possible is entailed by the theory. Which cases of non-uniformity are possible, and which are not, also follow from the theory.

Theories which characterize variation at the level of the entire grammar, or at the level of the individual construction, cannot provide for uniformity and non-uniformity (i.e. variation) in a way which is both empirically justified and theoretically interesting. Consider first the possibility that the choices of grammatical alternatives are made for the language as a whole by the setting of a general parameter. Such a theory predicts that every language should be uniform in the
sense characterized here. The grammar of English, which is uniform with respect to Negation, Wh-questions and VP-displacement, can be analyzed as [+DS] (for the past and present tenses). How then is Monnese analyzed? Monnese would have to be [+VR] and/or [−DS] because VR is grammatical in Negation. Yet it must also be [−VR] and/or [+DS], since VR is ungrammatical in Wh-questions, where DS is found instead. Since the general parameters cannot analyze the non-uniform system, it is necessary to posit more specific parameters, such as [+− VR to T] or [+− VR Negation] and [+− VR to C] or [+− VR Wh-question]. Monnese would have a positive value for the Negation parameter and a negative value for the Wh-question parameter. At this point the system is describing the attested combinations, and nothing more. Anti-Monnese is characterized just as easily as Monnese: it is [−VR to T] or [−VR Negation] and [+VR to C] or [+VR Wh-question]. Moreover, the ‘maximally non-uniform’ systems in (10), which are impossible in the OT-based theory presented here, are just as simple to represent as Monnese (and Anti-Monnese). They specify three different choices for Negation, Wh-questions and VP-displacement. While I have framed these observations in terms of positive and negative values for parameters, the logic of the argument holds for any theory of variation which relies on devices with the same critical properties. For example, positing features which have particular values (such as ‘strong’ and ‘weak’, ‘interpretable’ and ‘uninterpretable’) leads to the same uncontrollable characterization of variation, in the absence of a strong theory of their distribution and properties.

At this point the analysis of variation that is on the table provides a separate statement encoding the choice of form for each ‘construction’, along the lines just sketched. It is telling that the non-uniform distribution of DS in Monnese leads Culicover (2008, p. 32) to conclude that it is not a real case of DS but rather an instance of ‘do-periphrasis’, a separate construction. Non-uniformity, under the assumptions he is working with, necessitates proliferation of constructions. (Grimshaw, 2013, discusses this case in more detail.)

Any case-by-case approach to grammars contrasts sharply with the OT account offered here, in which the choice made in one of the candidate sets bears directly on the possible winners in the other candidate sets, because a single ranking of the constraints must generate all of the choices.

The danger of case-by-case hypotheses is that any combination of parameter settings or any combination of construction types should constitute a possible human linguistic system. This problem is tied, inevitably, to learnability problems. I noted above that every candidate set can provide information about the grammar, i.e. the ranking of the universal constraints in place for the language. Any ranking that is established on the basis of any candidate set holds for all candidate sets. In contrast, if parameters must be set, or properties fixed, not for the grammar as a whole but for certain sub-areas of the grammar, such as a construction or a set of constructions, then the specification must be learned separately for each sub-area. If properties of individual constructions must be learned, then each construction must be individually evaluated in order for learning to occur.
Along with the role of generalization in learning, depth of explanation is sacrificed under theories in which grammatical decisions are independent of each other. As emphasized in Grimshaw, 1997, in such theories a change in the characterization of one part of the grammar need have no consequences at all for the characterization of another part. Under the principles of OT, the analysis of any aspect of a language entails properties of the analyses of all languages. The analysis of any candidate set entails properties of the analysis of all candidate sets in the language. Both uniformity and non-uniformity within and across grammars are an inevitable consequence of the theory of candidate sets and constraint interaction.

References


Travis, L. 1984: *Parameters and Effects of Word Order Variation*. PhD Dissertation, MIT.