TYPOLOGICAL STRUCTURE AND PROPERTIES OF PROPERTY THEORY

By

NATALIE R. DELBUSSO

A dissertation submitted to the

School of Graduate Studies

Rutgers, the State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Linguistics

written under the direction of

Alan Prince and Bruce Tesar

and approved by

____

____

____

____

New Brunswick, New Jersey

January, 2018

ABSTRACT OF THE DISSERTATION

Typological Structure and Properties of Property Theory

By NATALIE R. DELBUSSO

Dissertation Directors:

Alan Prince and Bruce Tesar

Formal typological analysis provides an otherwise unobtainable level of insight into both theories and the linguistic facts they analyze. This dissertation develops Property Theory (Alber & Prince 2016, 2017, in prep., Alber, DelBusso & Prince 2016), a theory of typological structure in Optimality Theory (OT; Prince & Smolensky 1993/2004). The list of languages generated in an OT factorial typology shows what the theory predicts, but not why it does so nor how it organizes the languages in the typological space. Property analysis answers these questions, finding the core structure that emerges directly from the logic of OT.

As a theory of formal OT typologies, Property Theory has a complex internal structure. The dissertation develops algorithms to translate between the formal objects of Property Theory (properties) and those of OT (ranking conditions). It examines cross-property dependencies and sufficient conditions on a set of properties for it to generate OT grammars, and thus an OT typology.

In taking typologies themselves as objects of study, property analysis leads to a re-conception of core constraint relationships and identification of classes of intensionally equivalent systems that share an internal formal structure while differing in the empirical areas analyzed. The dissertation develops a typological definition of stringently-related constraints and shows that systems with such constraints have a common structure, explaining diverse data in the same way. It shows that this organization characterizes analyses deriving the Final-Over-Final Condition, a typology of possible cross-linguistic syntactic structures.

Acknowledgements

My deepest gratitude goes to my committee members: Alan Prince, Bruce Tesar, Jane Grimshaw, and Nazarré Merchant. Alan profoundly impacted my work since our first interactions, fundamentally changing my approach to research and pushing me to always go onward and upward; this dissertation could not exist without him. It has been a privilege to learn from and collaborate with him over the past several years. Bruce taught me so much about how and why to argue formally and to meticulously probe theories, especially one's own. I thank him too for crucial assistance in gaining computer-programming skills. Jane's work and way of thinking about research inspired me from the first day of grad school. She has an uncanny ability to see straight to the core of a proposal (and its weaknesses). Naz and his work greatly influenced mine, and far beyond an external committee member he has been a mentor and collaborator over the last few years. Feedback from and interactions with all four helped a collection of ideas become a dissertation.

Thanks to all my teachers and QP committee members in the Rutgers linguistics graduate program, especially Akin Akinlabi, Mark Baker, Maria Bittner, José Camacho, Veneeta Dayal, Viviane Déprez, Liliana Sanchez, and Roger Schwarzschild. I am grateful to the linguistics department for providing financial support throughout my graduate studies, and to Marilyn Reyes for help in administrative matters.

Thanks to collaborators, mentors, and friends outside of Rutgers: Birgit Alber, Eric Bakovic, Brett Hyde, and especially Will Bennett; and to department colleagues: Ümit Atlamaz, Nick Danis and Hazel Mitchley (to whom I'm also indebted for careful reading of various drafts and notes).

Finally, many thanks to my family and non-linguist friends for their support and encouragement over the course of my graduate career. And especially to my mom, Linda, whose support got me through the first day, to the last.

Table of Contents

Table of Abbreviations and Notations

∘	fusion
⊕	join
>	domination in ranking
$\rightarrow_C$	Order relation in EPO(C)
$\sim_C$	Equivalence relation in EPO(C)
α/β	Property values left-to-right/right-to-left
AOT	Abstract OT
BPP	Border Point Pair
C	Constraint
C(*q*)	Violations of candidate *q* assessed by C (C function from Cand to ℕ)
C[K]	C filtration of cset K (C function from 2^{Cand} to 2^{Cand})
COT	Concrete OT
cset	Candidate set
CT	Comparative tableau
EPO	Equivalence-augmented Privileged Order
ERC	Elementary Ranking Condition
FRed	Fusional Reduction algorithm
Γ	Grammar
h[ ]	OT filtration by h, a sequence of Cs ∈ CON
HB	Harmonically Bounded
JDG	Join disjunct grammars algorithm
κ	Constraint class

L	Language
L(ERC)	L-set of ERC
λ	linear extension of a grammar (leg)
MIB	Most Informative Basis (ERC set)
MOAT	Mother of all Tableaux
ns	Narrow scope
P	Property
PA	Property Analysis
PA(T)	Property Analysis of typology T
PvE	Property values to ERCs algorithm
PT	Property Theory
Σ(P)	Scope of P
SKB	Skeletal Basis (ERC set)
T	Typology
T_{OT}	OT typological partition
UVT	Unitary Violation Tableau
US	Universal Support
v	P value, $v \in \{\alpha, \beta\}$
VT	Violation Tableau
W(ERC)	W-set of ERC
ws	Wide scope

Table of OT Definitions

Key to references: ADP: Alber, DelBusso & Prince 2016; BP: Brasoveanu & Prince 2011; MP: Merchant & Prince 2016; P: Prince 2016b.

Term	*Definition*	*Reference*
Border point pair (BPP)	For typology T on a set of constraints Con_T, a pair of λs, (λ1, λ2) is a *border point pair* for two Γs, Γ1 & Γ2 if λ1 = $P\underline{XY}Q$, λ2 = $P\underline{YX}Q$, where P,Q are sequences of constraints in Con_T and X,Y ∈ Con_T, λ1 ∈ Γ1, λ2 ∈ Γ2.	MP: 81 (104)
Candidate (q)	An (input, output) pair and their correspondence.	P
Candidate set (cset, K)	All (input, output) pairs with the same input.	
Con_S	The set of all constraints of system S.	
Constraint (C)	A function from candidates to non-negative integers (violations).	
C Filtration, C[K]	C[K] = {q∈K|∄ z∈K such that C(z) < C(q)}: the set of candidates in K with the minimal value of C.	MP: 77 (99)
EPO(C)	For typology T, G = {Γ|Γ∈T}, EPO(C) = ⟨G, $<_C$, $\sim_C$⟩ (the order, $<_C$, and equivalence, $\sim_C$, structure of a C on Γs ∈ T).	MP: 84 (113)
Gen_S	Function defining the csets of system S.	P: 13
Grammar (Γ)	An ERC set delimiting a set of linear orders on Con_S that select the same language (set of optima).	P: 61
Harmonically Bounded (HB)	For Con_S and candidate set K, candidate q∈K is harmonically bounded in K iff ∀λ, ∃z∈K, z≠q: λ[{q, z}]=z.	BP; MP: 85 (116)

Language (L)	The set of optima under a given linear order on $\textsc{Con}_S$. Extensional: set of linguistic structures and mappings.	P: 61
MOAT(T)	For typology T, MOAT(T) = {EPO(C)| C∈$\textsc{Con}_T$}.	MP: 84 (114)
Optimality (filtration by hierarchy)	For H = C1>C2>…>C*n*, C*k*∈$\textsc{Con}_S$: {q∈K| q∈H[K]}. The set of candidates surviving sequential filtration by Cs ∈ H.	P: 27
Permuto-hedron	Geometric structure on the set of total orders of *n* elements, where each total order is a vertex connected to all total orders differing in a single adjacent transposition.	MP: §6.1
System (S)	<GEN, CON>	P: 13
Typohedron	Geometric structure of a typology where each Γ is a vertex connected to all other Γs with which it has a Border Point.	MP: §6.2
Typology (T)	1. For a set of constraints, $\textsc{Con}_T$, a partition of the set of orders on $\textsc{Con}_T$ is a typology iff there is a UVT, U, with columns corresponding 1:1 to Cs ∈ $\textsc{Con}_T$ and rows corresponding 1:1 to Γs ∈ T, such that each block of the partition is the ranking Γ of a row in U. 2. The set of all grammars of S (intensional); the set of all languages of S (extensional).	MP: 17 (12) MP: 9
Universal Support (US)	A set of csets ⊆ GEN necessary and sufficient to define all languages of the typology.	ADP

Unitary Violation Tableau (UVT)	A VT where each row gives rise to a distinct Γ.	MP: 16 (8)

1 Introduction

1.1 *Introduction*

Linguistic theories define a space of possible grammars, predicting the extent and limitations of variation between human languages—the *typology*. The grammars it generates instantiate distinct combinations of choices along the possible dimensions of variation. Knowing the predicted typology is crucial for assessing any hypothesis, often evaluated by comparison to the attested empirical one. However, knowing *what* the theory predicts is insufficient without understanding *why* it does so: how do the assumptions of the theory give rise to the predicted languages, and how does it explain and classify them? Answering these questions requires study of the internal structure of the typological space, analyzing the formal factors grouping and distinguishing the grammars.

This dissertation analyzes the formal structure of a specific concept of typological organization: Property Theory (PT; Alber & Prince (A&P) 2016a, in prep., Alber, DelBusso & Prince 2016). The results build on A&P's founding work to advance two central goals: formal development of the theory and its usability, and demonstration of the results of these advances in explicating the structure of typologies in Optimality Theory (OT; Prince & Smolensky 1993/2004).

Typological analysis is inherent to OT due to the centrality of factorial typologies. An OT factorial typology of a given system S, T_S, is all possible permutations (rankings) of a set of universal constraints on linguistic forms, CON, that give rise to distinct sets of optima (languages). While all permutations of CON, |CON|!, are possible ranking hierarchies, in many typologies several hierarchies result in the same extensional language; not all constraints conflict and are crucially ranked in all grammars. A *property analysis*, PA, discerns the crucial rankings that classify a typology: those necessary and sufficient to define every grammar. PT explicates the link between these *intensional* rankings and the *extensional* traits exhibited in the languages they generate (Alber, DelBusso & Prince 2016 (ADP)). A PA explains why the predicted typology arises from elements of the theory—GEN and CON—and its non-arbitrary and often non-obvious structure.

A PA analyzes a typology, T, into a set of *properties*, Ps, that antagonize sets of constraints in CON, X <> Y. A P generates two mutually exclusive *values* that rank the antagonists in opposing ways: α. X > Y, β. Y > X. Grammars are classified according to which value is instantiated in their rankings. Values often align with particular linguistic traits being optimal. For example, in the Elementary Syllable structure typology, EST (Prince & Smolensky, Prince 2016a, Merchant & Prince, A&P 2016b), languages differ in whether they allow onsetless syllables in optima. The extensional characteristic aligns with the grammar's value of a property antagonizing a markedness constraint, m.Ons—violated by syllables lacking onsets—with one of the faithfulness constraints {f.max, f.dep}—violated by deletion or insertion of segments, respectively. Each property defines a binary partition, and the entire typology is defined by a collection of properties: it is the partition resulting from their consistent value combinations. Within this space, grammars are grouped together and distinguished based on shared and unshared values.

PT explicates how the objects of the theory—the constraints and their interactions—generate languages. Variation as binary choice is a common theme in linguistic theory, from the idea of parameters in Principles and Parameters, where languages choose settings of parameters and the combinations thereof define a typology. Recent work in parametric theory also more explicitly seeks to understand the internal structure of typologies (Baker 2001, Roberts 2010, et seq., chapter 4 herein).

Work in PT has produced significant results, both in understanding typological organization and in solving fundamental problems in OT (Alber 2015a,b, A&P 2017, Bennett 2016, Bennett & DelBusso 2017, to appear, Bennett, DelBusso & Iacoponi 2016, Danis 2014, DelBusso 2016, McManus 2016). Alber, DelBusso & Prince (2016) use properties to prove a Universal Support (US), a set of candidate sets, csets, necessary and sufficient to generate all grammars, exemplifying the method for the stress system nGX (A&P 2017). Alber (2015a,b) develops a property-based theory of grammatical variation and diachronic change (used in the present chapter 4). Bennett & DelBusso (to appear) explicate the typological effects of systematic changes to constraint definitions in a set of Agreement-by-Correspondence (ABC) systems through PAs. They align the resulting properties with specific extensional predictions, defining the formal factors that generate the linguistic patterns. Bennett & DelBusso (in prep.) further analyze different definitions of GEN, including those lacking correspondence. PT analyses of the systems show what aspects of a theory are crucial to deriving consonant harmony and dissimilation, and how these can be instantiated in various ways, with divergent assumptions.

Data for the dissertation come from a database of analyzed typologies, both Concrete and Abstract OT systems. A Concrete OT (COT) system, S, analyzes some particular linguistic phenomenon, defining GEN_S, the set of allowable structures, and CON_S, the set of constraints assessing them (Merchant & Prince 2016/to appear (M&P)). Abstract OT (AOT) systems start with a set of constraint filtration profiles and examines the typology resulting from them. Though not intended to analyze a particular linguistic fact, study of AOT systems gives rise to significant formal results (as in M&P). They allow for generalization across extensionally diverse systems that share an *intensional* structure, and distill central interactions that may be obscured in larger systems. Analysis of AOT systems feeds understanding of COT systems, and some AOT systems have exact or near-exact COT correlates.

The results of the dissertation are embedded within the extensive formal development of OT (especially Prince 2002, 2016a,b, M&P) and PT (A&P 2016a, in prep.). It uses terms and definitions of modern OT and assumes basic familiarity with Entailed Ranking Conditions (ERCs) and their logic (Prince 2002, Brasoveanu & Prince 2011). All other terms and abbreviations are defined on first use and key OT terms are included in the glossary for reference.

1.2 *OT Typologies*

As PT is a formal theory of the structure of OT typologies, its formal development requires understanding of such objects. The formal structure of OT typologies is well understood, due especially to the results of Merchant & Prince (2016/to appear), reviewed in this section.

Extensionally, a typology is the set of languages of a system, S, where each language is a set of optima. *Intensionally*, it is a partition of the set of total orders over $\text{C}{\small\text{ON}}_S$, the set of constraints of a system S, where each part of the partition is a grammar (Γ). OT grammars are antimatroids, delineated by a set of rankings (ERCs) (M&P p. 9, Merchant & Riggle 2016). This ERC set may describe a single total linear order or a set of such orders. Each such order is a *l*inear *e*xtension of a *g*rammar, a *leg*, λ.

The core analytical objects are defined in (1); following M&P (p. 9), *grammars* are distinguished from *languages*. The former is a set of intensional rankings characterized by an ERC set; the latter is the set of extensional forms that are optimal under those rankings. Typological analysis can occur at both levels: *extensionally*, examining the list of languages generated, and *intensionally*, studying the rankings generating them.

1) *Definitions: Language, Grammar, Typology*

   a. *Language (L)*: the set of optima under a given constraint hierarchy.

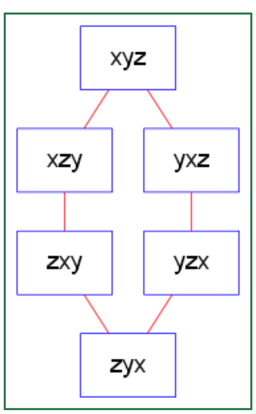
   b. *Grammar (Γ)*: an ERC set delineating a set of linear orders, λs, on CON that select the same set of optima (a language).

   c. *Typology (T)*: *extensional*: the languages of the system.

      intensional: the grammars of the system.

There is a natural geometry on the set of total orders, represented by a graph called a *permutohedron* in which each total order is a vertex connected to those from which it differs by a single adjacent transposition of two elements (M&P §6.1). For a set of n elements, the permutohedron is an $n-1$ dimensional object. The permutohedron for a 3-constraint system is shown in (2); it is a 2-dimensional hexagon with six vertices.

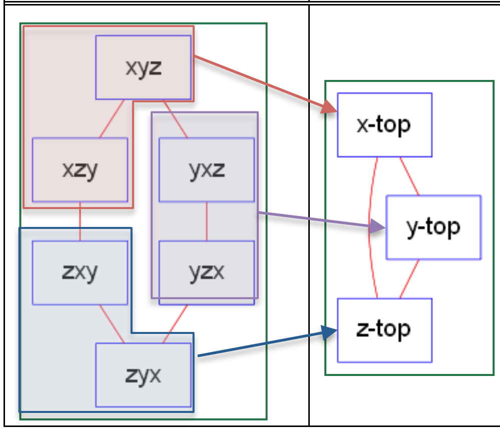
2) *3C permutohedron*



A typology maps to a permutohedron of the total orders over Con$_S$, with Γs represented as connected regions, sets of adjacent total orders. The AOT system called the 'tops' in M&P (p.169) and T.1|2 in DelBusso & Prince (in prep.), has three Γs, each defined by a ranking in which a single constraint dominates the other two, which are not themselves crucially ordered. When all total orders that lie within the same grammar are collapsed into a single node the resulting object is a *typohedron* (M&P §6.2). In T.1|2, each Γ covers two adjacent vertices of the 3C permutohedron, producing the typohedron in (3).

3) *T.1|2 typohedron*

Γs on permutohedron	*Typohedron*



Adjacency between two grammars Γ_1 and Γ_2 in a typohedron is defined by a *border point pair* (BPP): a pair of λs differing in the single adjacent transposition of two constraints, that belong to different Γs ((4), from M&P:81 (104)).

4) *Def. Border Point Pair (BPP)*. For a typology T on a set of constraints Con$_T$, a pair of λs, (λ_1, λ_2) is a *border point pair* for two Γs, Γ_1 and Γ_2 iff $\lambda_1 = P\underline{XY}Q$ and $\lambda_2 = P\underline{YX}Q$, with P,Q sequences of constraints from Con$_T$, X,Y ∈ Con$_T$, $\lambda_1 \in \Gamma_1$ and $\lambda_2 \in \Gamma_2$.

For example, in T.1|2, λs, λ_1 = $\underline{xy}z$ and λ_2 = $\underline{yx}z$ are a BPP for Γs x-top and y-top: the two legs differ in the adjacent transposition of x and y, and one belong to each Γ (P is empty, Q = z).

A partition of a permutohedron where all parts are defined by ERC sets is a *grammatical* partition. OT typologies are a subclass of such partitions, proven by M&P (Theorem (189)) to be those that can be represented with a Unitary Violation Tableau (UVT; Prince 2016a, (5) from M&P:16 (8)), or, equivalently, an acyclic *MOAT* ((6) from M&P:17 (12)).

5) *Def. Unitary Violation Tableau* (UVT). A violation tableau in which each row gives rise to a distinct grammar.

6) *Def. OT Typology* (T_{OT}). A partition of the set of orders on a set of constraints CON_S is a typology iff there is a UVT U, with columns corresponding 1:1 to the constraints ∈ CON_S and rows corresponding 1:1 to the Γs ∈ T, such that each block in the partition T is the ranking grammar of a row in U.

Each C in a T *filters* the candidate set, assigning a non-negative value to each candidate. The set of candidates with the minimal value assigned are those that pass through its filtration, *survivors* of C; all others are *rejected*. As with a constraint, so with an ordered sequence of Cs, a hierarchy, h: each C in h successively filters the candidates surviving the preceding Cs (M&P p. 77ff). A hierarchy is *decisive* if is determines a violation-profile unique optimum (co-optima have the same violation profile).

7) *Filtrations*

For a set of candidates, K:

a. $C[K] = \{k \in K: \nexists q \in K, C(q) < C(k)\}$ (M&P p. 77 (99))

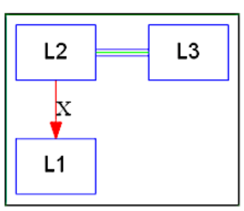
b. h[K]: for h = an ordered sequence of Cs ∈ Con, (X, Y, Z), h[K] = Z[Y[X[K]]]

c. Def. Decisive hierarchy: a hierarchy h is *decisive* for K if |h[K]| = 1.

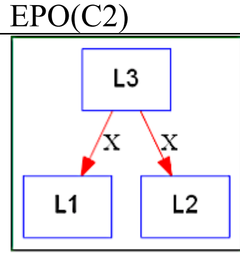
A MOAT (Mother of all tableaux) is a collection of Equivalence-augmented Privileged Orders (EPOs) that encodes the filtration patterns of each constraint in CON over the set of Γs ∈ T as an order and equivalence structure (M&P). As an example, the MOAT of the simple stringency system T_{2Core} (chapter 3) is shown in (8). Order relations are indicated by arrows, labeled for the other C in the BPP giving rise to the arrow. Double blue lines represent equivalence; the connected grammars are in an equivalence class. In this system, C2 and X order pairs (L1, L3) and (L2, L3) differently; C1 and X order (L1, L2) differently; L2 and L3 are equivalent for C1; L1 and L2 are non-comparable for C2.

8) *T_{2Core} MOAT*

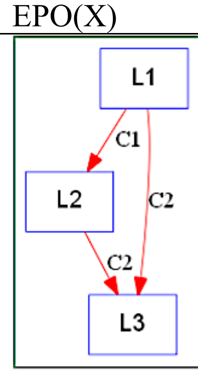
EPO(C1)



EPO(C2)



EPO(X)



While there are many possible UVTs for a given T, there is a single MOAT: EPOs record the two relations that matter in optimization, order and equivalence, but no specific violation values, as many different ones produce the same filtration patterns (M&P). The present work defines a *minimal UVT* (mUVT) as a UVT that derives from the MOAT and uses the minimal possible violation values.

9) *Def. minimal UVT* (*mUVT*): a UVT where for each Γ in a row of U, C(Γ) is length of longest arrow chain separating Γ's equivalence class from the top equivalence class of EPO(C).

M&P prove that the MOAT fully determines every Γ in T, allowing for argumentation from MOAT properties to typological properties. In chapter 3 of this dissertation, MOAT structure is used to identify the constraint relationships—conflict, stringency, and equivalence—that exist in T, deduced from comparing EPO structures.

With the definition of a T, formal relations between Ts can also be described. Two Ts, T1 and T2, are equivalent if their MOATs are isomorphic: the grammars of each have equivalent rankings, defining a bijection between the CONs (M&P §0.3.1). Intensional typological equivalence is an underlying theme running throughout this dissertation. Analysis at the intensional level draws out the structural commonalities between systems of diverse phenomena, in distinct areas, thus allowing for broader generalizations about the organization of linguistic systems. Chapter 3 analyzes the shared structure of systems with stringency constraints. The results allow for understanding of an entire class of typologies, which explain the distribution of different extensional traits in parallel intensional ways. Chapter 4 analyzes a set of syntactic typologies in three related ways, and shows that all resulting typologies are intensionally equivalent.

1.3 *Property Theory and Analysis*

A *property analysis of a typology T*, PA(T), analyzes the intensional rankings structuring the typology, finding the grammatical choices that define the system. A PA delineates these rankings and aligns them with extensional linguistic structures, *traits*, showing how the formal choices relate to the predicted languages. The properties are the intensional dimensions along which the system is organized.

This section provides an overview of the core mechanisms of PT, drawing especially on Alber & Prince (2016a). See also Alber & Prince (2017) and Alber, DelBusso & Prince (2016) for introductions to the central concepts.

1.3.1 Properties

A *Property Analysis* (PA) contains a set of *properties*, Ps. Ps are stated in the form X <> Y, where X and Y are the P *antagonists*, and the *values*, α and β, are the mutually exclusive rankings generated by reading the ranking relation in either direction: α. X > Y and β. Y > X (A&P 2016a). Each value generates an ERC set (chapter 2 develops methods for converting a ranking statement to ERCs), partitioning the set of total orders in a T. A Γ in a T has a value, P.α or P.β, when it non-trivially entails the ERCs of that value and thus contradicts the other[1]. A PA is a set of Ps that define all and only the Γs ∈ T as the possible distinct combinations of values.

In the most basic case, X and Y are single constraints, Cs. The P values are their two possible orderings, generating ERCs with a single W and L sets. Some Ts can be completely analyzed with such Ps; a total-order T, where each Γ is a single λ, is an example. However, in a given system, a pair of Cs may not conflict in all or any Γs. For every T, there is a defining set of crucial constraint conflicts. Groups of constraints can act together as a *class* in an antagonist so that conflict is between sets rather than individual pairs.

---

[1] When a Γ is consistent with both values, P is *moot*; see below.
[2]This scope can be stated positively with disjunction if the P1 moot region is definable by a value set.
[1]This chapter is also indebted to Nazarré Merchant for input and assistance, especially in the

X and Y abbreviate *classes* of constraints, κs (A&P 2016a, in prep.; chapter 2 herein). Properties involving classes recognize a level of shared ranking information that may not be representable in a single ERC: a set of Γs shares a ranking where *some* C in a κ is ranked relative to the antagonist, but the individual Γs may differ in which particular C that is. A specific C is determined by an operator (op), *dom* or *sub*, functions that return the extremes in a linear order, λ, on Cs in κ, the highest and lowest ranked, respectively (A&P 2016a:§II). For example, (10) shows the C returned by each op for two linear orders. If each λ is in a different Γ, then the Γs share that κ.op is ordered relative to the antagonist but differ in the C in κ.

10) *Dom & Sub operators*

κ = {x, y, z}	λ1: xyz	λ2: zxy
κ.*dom* = {x, y, z}.dom	x	z
κ.*sub* = {x, y, z}.sub	z	y

The ops have quantificational force by virtue of referring to the extremes of a total order (A&P 2016a). If *any* C in a κ dominates x, then the highest, κ.dom, does, transitively. Conversely, if x dominates κ.dom, then it dominates *all*. Dom is equivalent to Boolean disjunction when dominant and conjunction when subordinate. The ERCs generated by such values have multi-W/L-sets, representing this dominator disjunction/subordinate conjunction. Sub is the reverse: if κ.sub, the lowest member, dominates x then *all* Cs ∈ κ do (conjunction); but if x dominates *any* κ C, then it dominates the lowest, κ.sub (disjunction). As subordinate disjunction is not ERC representable, values with a subordinate κ.sub generate disjunctive ERC sets, with each set having a different L. A value with a dominant κ.sub is a conjunctive ERC set, where ERCs share L-sets and differ in Ws, as exemplified in (11) for a 2C κ.op, {yz}.op.

11) *Ops and ERCs*

a. *Dom*: P: x <> {yz}.dom	b. *Sub*: P: x <> {y,z}.sub
α. x > y *and* x > z `WLL` β. y >x *or* z > x `LWW`	α. x > y *or* x > z `WLe|WeL` β. y > x *and* z > x `LWe,LeW`

A κ can be a singleton, in which case op is omissible, either returning the same single C.

To give an example, in the typology of the basic syllable system EST (Prince & Smolensky 1993/2004, M&P, A&P 2016b), the faithfulness constraints f.max and f.dep are a κ in the PA. Both of the markedness constraints, m.Ons and m.NoC, are (separately) ranked relative to the *sub* of this κ (12).The values align with the extensional traits of onsetlessness and coda allowedness, respectively. Under each of the α values, one of the Cs in the κ is dominated; which is determined by the value of P3, which orders these. Values of P3 correlate with whether insertion or deletion of segments is optimal in unfaithful mappings.

12) *PA(EST): Properties*
C order in ERCs: m.Ons-m.NoC-f.max-f.dep

P	*Values*	*Extensional trait*
P1: m.Ons <> {f.max, f.dep}.sub	α. `WeLe	WeeL`	onsets required
	β. `LeWe, LeeW`	onsetlessness allowed
P2: m.NoC <> {f.max, f.dep}.sub	α. `eWLe	eWeL`	no codas
	β. `eLWe, eLeW`	codas allowed
P3: f.max <> f.dep	α. `eeWL`	insertion
	β. `eeLW`	deletion

1.3.2 Scope: property interdependencies

Work in PT has shown how some rankings are dependent on others (A&P). This has correlates with extensional traits: some choices of linguistic structure are contingent on others. For example, only among languages that allow codas in syllables is there a choice between allowing complex codas versus only single consonants.

Intensionally, for a given P, some Γ(s) may have neither value if both are consistent with the ERC set, in which case P is *moot* (A&P). Such Ps only distinguish among a subset of the Γs. While all Ps are binary partitions of the entire set of λs in T, in that every total ordering, λ, satisfies one value or the other, they may not be such partitions of the set of *Γs*, because a Γ can include λs in both parts of P's partition. In such a Γ, the P antagonists are not crucially ranked, occurring in either order in some λs. A P *scope*, Σ(P) defines its domain; for a *wide-scope* P (wsP), all Γs have a value but for a *narrow-scope* P (nsP) some do not. These are defined by positive Boolean combinations of other P values (A&P 2016a:10).

13) *Def. Scope*. For a PA = {P1,…P*n*}, the scope of a P1 ∈ PA, Σ(P1), is the subset of Γs ∈ T, {Γ1,…,Γ*m*}, that have a value of P1.

   a. *Wide scope* (ws): {Γ1,…,Γ*m*} = T.

   b. *Narrow scope* (ns): {Γ1,…,Γ*m*} ⊂ T & is defined by a positive Boolean combination of P values from other Ps ∈ PA, {P2,…P*n*}.

P value scope definitions pick out the set of Γs sharing that value description, and possibly differing in other values. These can be single values, value conjunctions, disjunctions, and combinations thereof. Negative scopes, such as ¬P1.α, are illicit. If P1 is ws, this is equivalent to a single-value scope for the opposite value (P1.β); but if ns, ¬P1.α includes both P1.β and Γs for which P1 is moot[2]. Cyclic scopes, where the scopes

[2]This scope can be stated positively with disjunction if the P1 moot region is definable by a value set.

of a set of Ps are co-dependently defined by each other, are excluded by virtue of being undefined (a loop).

Conjunctive and disjunctive scopes both arise frequently in analysis. Conjunctive scopes occur when multiple rankings are jointly necessary for antagonists to conflict; examples in COT systems include Danis (2014) and some of the syllable systems analyzed by DelBusso & Prince (2015). Disjunctive scopes occur when there are multiple other rankings, each independently creating the conditions for such conflict; examples arise in systems with antagonized sets of stringency constraints (chapter 3 and Alber 2015ab), and in systems with overlapping subPAs (Bennett & DelBusso to appear).

A full PA defines both the set of Ps and their scopes, producing the set of (potential) Γs that corresponds to all consistent value combinations given the scopes. A PA that generates T, defining exactly its set of Γs, is a *valid* PA, the central topic of chapter 2. A full PA is represented in two ways: in a value table and a treeoid. A value table lists Ps as columns and the possible combinations of their values as the rows. A treeoid is a directed acyclic tree graph augmented by various kinds of lines (A&P 2016a:11, 2017). Ps label nodes of the treeoid and are connected to their value nodes by double red lines, indicating a mutually-exclusive choice. The value nodes dominate any P(s) whose scope(s) they define, represented as single blue lines. Dotted blue lines indicate disjunctive scope in the sense that for a P dominated by dotted lines any Γ having a value of *any* of the dominating nodes has a value of P.

1.4 *Dissertation outline*

The core dissertation chapters focus on three major areas: formal development of PT and conditions under which a set of Ps generates a T (chapter 2); understanding the typological structure of classes of OT systems, specifically those with stringently-related Cs (chapter 3); and, the PT explanation of typological organization, as exemplified by a word order typology, compared to a recent parametric theory (chapter 4).

1.4.1 Valid Property Analyses

A PA analyzes a particular typology, T. Of a set of Ps, two questions can be asked: first, does it generate all and only the Γs of a given system under analysis, S? and second, does it generate any OT T? Chapter 2 defines two kinds of *valid PAs*, aligning with these two questions. It then examines the conditions for a set of Ps to be a valid PA in the second, more abstract sense, generating an OT partition of a set of λs.

Since the objects of PT—sets of property values—and OT—ERC sets—are distinct, formally precise methods are needed to convert one to the other. Chapter 2 presents algorithms to calculate predicted grammars from a set of values in a property analysis, building on DelBusso & Merchant (in prep.). It proposes the Join-Disjunct-Grammars (JDG) algorithm, which is used in assessing PA validity by determining if the value sets generate non-overlapping Γs. The algorithms provide computational analytical tools that facilitate Property Analysis automation, and are incorporated into OTWorkplace (Prince, Merchant & Tesar 2007-2017), a software package for rigorous OT analysis. Automations ensure accuracy and extend the reach and utility of the theory, allowing for analysis of large and complex systems where manual approaches are untenable.

The chapter also defines a relationship between properties within a PA, formalized in the concept of a *resolver P,* resP. A resP is a property that antagonizes the Cs in a class, κ, in another P. Such Ps are shown to establish sufficient conditions for a set of Ps to generate a grammatical partition.

1.4.2 The structure of stringency systems

Analyses of a wide range of systems identify intensional structural equivalences across systems modeling distinct extensional phenomena. The typological perspective also gives rise to a reconceptualization of core constraint relationships. Typological structure depends on the set of filtration patterns of the Cs ∈ CON, which may be the same across systems in which different Cs evaluate different structures. OT systems can be understood and classified according to kinds of C relationships that occur therein, discovered through property analysis. Such relations are not limited to conflict. Chapter 3 is a detailed analysis of the structure of typologies involving constraints in a *stringency* relationship, a relation of non-conflict.

Stringently-related Cs are common in OT analysis, used to derive implicational universals: if a language has trait *x*, it has trait *y*, but not vice versa. The chapter develops a new formal definition of stringency inherently linked to OT typological structure by referring to *filtration patterns* rather than violation counts or C definitions. Two Cs may appear to be in a stringency relationship based on their definitions, but fail to behave as such within a given system due to other factors, such as GEN. Filtration stringency is identifiable from a MOAT. This leads to a further identification of a relation of *partial stringency*, where Cs stand in the relation over only some but not all Γs in T. The chapter further classifies the MOAT and property correlates of other constraint relations, conflict and equality.

Detailed development of the PA structure stringency system shows the core set of interactions that occur in all systems with such Cs: a typology of segmental faithfulness (Alber 2015ab, chapter 3) is intensionally identical to one of syntactic structure (chapter

4), despite non-comparable extensional languages. Expanding the basic system in systematic ways refines or iterates the structure. The same properties also occur in the PAs of partial stringency, though in the context of other Ps. These Ps align with extensional traits of linguistic scales, characterizing how a scale manifests in a language.

Analyzing the core relations allows for understanding of a range of phenomena and provides analytical tools. When stringency relations are identified, a set of properties can be immediately stated, yielding complete understanding of some simple systems and providing a hook into the structure of more complex cases.

1.4.3 The Final-over-final condition and typological structure

While previous chapters focus on formal aspects of PT and intensionally equivalent classes of systems, chapter 4 analyzes the PAs explanation of both a specific system and typological organization more generally. This is compared to a recent proposal for the structure of syntactic typologies in parametric theory: Parameter Hierarchies (Reconsidering Comparative Syntax project (ReCoS), Roberts 2010, 2012, et seq.).

In the theory of Parameter Hierarchies (PH), *parameters* and their settings define the dimensions of variation, and the predicted typology is the possible combinations of settings. Parameters are organized in a common hierarchical structure, resulting in a fixed set of ordered choices among their settings. There is an intuitive conceptual similarity between parameters, settings, and hierarchies in parametric theories, and properties, values, and treeoids in PT.

The chapter develops a set of analyses of the Final-Over-Final Condition (FOFC, Biberauer et al. 2014), a cross-linguistic generalization of possible word orders and a systematic gap therein. The analyses use sets of stringency constraints; their intensional and extensional equivalences show the essential components required to derive the condition. These are compared with Biberauer et al.'s (2014) analysis. The PA structure closely resembles the Parameter Hierarchy of the FOFC typology, but diverges in ways that show deeper differences between the theories. In PT, typological structure follows directly from the objects and logic of OT itself: the constraints and their conflicts over a set of candidates that define the Γs.

2 Valid Property Analyses

2.1 *Introduction*

A Property Analysis of a typology of a system S, PA(T_S), analyzes T_S into a set of properties whose values generate all and only the grammars (Γs) of T_S. A central, more general question in Property Theory is the conditions under which a set of properties, Ps, generates *any* OT typology at all. As Merchant & Prince (2016/to appear) show, OT typologies are a certain class of partitions of the set of total orders of constraints in CON—those having an acyclic MOAT/UVT. A given set of Ps is not guaranteed to yield such an object. This chapter examines conditions on set of Ps to be a *valid PA(T)* in this sense. It is deeply embedded in and indebted to the extensive development of Property Theory (Alber & Prince 2016a, in prep. (A&P), DelBusso & Merchant in prep.), and on OT typologies in Merchant & Prince (2016/to appear; M&P), and building on concepts in these works.[1]

The results rest on 1) having formally explicit methods to translate between property value sets and ERC grammars (Γs); and 2) the notion of a *resP*. The first, left implicit in previous work, is complicated by the fact that P antagonists are often not single constraints, Cs, generating ERCs with a single W and L, but *constraint classes*. These were developed by A&P (2016a, in prep.), further analyzed in DelBusso & Prince (in prep.; D&P) as binary hierarchical tree structures over C sets. The present chapter refines and formalizes that conception, generalizing to (non-binary) trees (§2.3).

[1]This chapter is also indebted to Nazarré Merchant for input and assistance, especially in the formalization of classes and class trees, and in the proofs in §2.5. Merchant wrote code for the JDG algorithm in the PA checker function of OTWorkplace (Prince, Merchant & Tesar 2007-2017).

As discussed in chapter 1, a specific C in a class is designated in a given total order by the operators *dom* and *sub*. These operators result in complex ranking conditions that, in the case of sub, are not representable by an ERC set. Both operators generate disjunctions in ERCs, but they crucially differ in whether the disjunction is grammatical. A dominant κ.dom in a value, P.v, κ.dom>x, generates single ERCs with dominator (W) disjunction, a possible OT grammar, Γ.[2] However, a subordinated κ.sub, x>κ.sub, has subordinate (L) disjunction, and any such property value generates a disjunction of ERC sets, with each disjunct differing in the specific C(s) dominated. As A&P have shown, such a value does not define a Γ.

This chapter presents two algorithms to generate ERC sets from the PA value sets. The P-values-to-ERCs algorithm (PvE; §2.4.1, DelBusso & Merchant in prep.) converts a P value, a statement of a ranking condition, to (sets of) ERC sets. This formalizes a step implicitly assumed in work PT, allowing for automatization. The chapter then introduces the Join-Disjunct-Grammars algorithm (JDG; §2.4.2), which takes a set of values and returns the ERC set it describes. JDG uses core elements of OT logic: the Fusional Reduction algorithm (Brasoveau & Prince 2011) and join operator (Merchant 2008, 2011), from which it draws its name.

JDG provides a solution to the issue of generating a single OT Γ from a PA value set that includes disjunctive values and is central to assessing PA validity. For a valid PA, each value set defined by the PA must result in a *conservative* output of JDG. Merchant (2008, 2011) shows that a join is conservative when it is equal to the union of the ERC sets joined, excluding any additional total orders that are not in any of these sets. In §2.5,

[2]Recall that a Γ_{OT} is an antimatroid described by an ERC set, not necessarily a total or partial order (Merchant & Riggle 2016).

the conditions for JDG conservativity are examined. The presence of Ps with values that generate disjunctive ERC sets results in cross-property dependencies; the PA must include some other Ps in order for JDG to be conservative.

The needed Ps are examined in §2.6, which introduces the concept of *resP*, defined as Ps that draw their antagonists from a constraint class in another P, and antagonize the Cs within this class. These are argued in §2.7 to be sufficient for conservative JDG outputs, leading to a Theorem of sufficient conditions for a set of Ps to generate a grammatical partition, crucial for validity.

2.2 *Valid Property Analyses*

A *Property Analysis*, PA, contains a set of *properties*, Ps (1), defining two opposing ranking conditions (A&P 2016a).

1) *Def. Property*, P: antagonized constraints classes, κα.op <> κβ.op, with values, α: κα.op > κβ.op and β: κβ.op > κα.op, generating ERC sets defining mutually exclusive rankings that partition the set of λs.

A P is a *binary* partition of the set of total orders, λs: every λ satisfies the ERCs generated by one value and is inconsistent with the other. Neither value can be empty because a value describes a ranking between Cs. Since a typology is a partition of the set of all possible linear orders, some orders instantiate one ranking and some the other. In the context of a given T, some Γs may include λs in both halves. In this case, P is moot in these Γs, as the ERC set of either value if consistent with Γ.

A *valid PA* analyzes a typology, generating grammars. Of a set of Ps, two questions can be asked: 1) does is generate the typology of a given system S, T_S? and, 2) does it generate any OT typology? This correlates with M&P's two definitions of a typology as:

1) as the collection of grammars of a system (p. 9); and 2) a subset of partitions of the set of total orders over CON that has a UVT/acyclic MOAT (p. 17 (12), definition repeated from chapter 1 in (2)).

2) *Def. OT Typology* (T_{OT}). A partition of the set of orders on a set of constraints Con_S is a typology iff there is a UVT U, with columns corresponding 1:1 to the constraints $\in Con_S$ and rows corresponding 1:1 to the Γs ∈ T, such that each block in the partition T is the ranking grammar of a row in U.

Corresponding, there are two concepts of a *valid PA*. A system-specific *valid PA(T_S)* is a PA that generates the typology of particular system S, following the definition of Alber & Prince (2016a:1) in (3).

3) *Def: Valid PA(T_S):* a PA(T_S) of a typology T = {Γ1,…Γ*n*} is a set of properties, {P1,…,P*n*}, such that each allowed, logically consistent choice of values yields a Γ∈T and each Γ∈T is so described.

In a valid PA(T_S), there is an bijection between the possible value sets of the Ps and the Γs ∈ T. In many cases, there are multiple valid PAs of a given T. D&P show this in detail for Weak Order Typologies (WOTs) and work in PT—including the present text—broadly demonstrates it. Determining if a set of properties is an analysis of S requires calculating the ERC sets resulting from each P value, using the algorithms developed in §2.4, and checking whether there is such a bijection. This validation can be done with the Property Analysis checker in OTWorkplace (Prince, Merchant & Tesar 2007-2017).

The second concept of a valid PA is more abstract: whether a set of Ps, and the sets of their value combinations, generates any OT partition. Each possible value combination is the P characterization of a grammar; since these are not guaranteed to be OT Γs, the notation pΓ is used (4); each pΓ is a row of a PA value table.

4) *Def. pΓ:* Given a set of PA, {P1,…,P*n*}, a pΓ is a unique set of scopally allowed consistent P values of Ps ∈ PA.

The definition of a *valid PA(T_{OT})* is a P set that describes an OT partition, absent a particular system under analysis. To be so, two conditions must be met: a) each possible value set, pΓ, generates an OT Γ, an ERC set defining a set of λs; and b) the Γs can co-exists in an OT partition.

5) *Def*: *Valid PA(T_{OT})*: A *valid PA(T_{OT})* is a set of Ps, {P1,…,P*n*} and the set of their possible value combinations, {pΓ1,…,pΓ*m*}, s.t.:

   a. Each pΓ generates an OT Γ, an ERC set that delineates a set of λs, total orders over CON;

   b. The set of Γs is an OT typological partition of the set of permutations of total orders over CON (a partition with a UVT).

Meeting the conditions for a valid PA(T_S) entails meeting those for a PA(T_{OT}), since S is, by assumption, a T and so an OT partition. Conversely failing to be a PA(T_{OT}) entails failing to be a PA(T_S) for any S, as the failed PA does not describe any OT T. However, it is possible for a set of Ps to describe a T_{OT}, meeting (5), but fail to generate a given system.

To satisfy the first condition, (5)a, all value sets, pΓs, must generate OT Γs. A pΓ is a set of ranking condition statements. Each value must be converted to a (set of) ERC sets, using the PvE algorithm (24), and then the entire set of values into a single ERC set. As Ps can generate disjunctive ranking conditions, the individual value ERCs cannot simply be amassed. The Join-Disjunct-Grammar algorithm, JDG (26) uses Merchant's (2008) join operator to produce a non-disjunctive ERC set for a pΓ.

To satisfy the second condition, (5)b, the set of pΓs joined be an OT partition. As a partition, pΓs are necessarily disjoint value sets, with none defined by a superset of the values defining another. An OT partition is one having an acyclic MOAT or, equivalently, a UVT (M&P). It is sometimes possible for a grammatical partition to describe a Harmonic Serialism typology, T_{HS}; for example, M&P's single split bot (§5.2.1) is a possible T_{HS}, though it cannot be a T_{OT}.

Failure to meet these criteria can result from the presence of Ps that generate disjunctive rankings. These arise from the presence of constraint classes, κs, with the operator sub in the P antagonists.

2.3 *Constraint classes*

C-classes are central in PT. As A&P establish and much subsequent work show, P antagonists are not always single Cs, but sets of Cs (sets). Such classes recognize a higher level of grammatical similarity, where grammars share not that a specific C is dominated, but that one in a set is. For example, in the PA of basic syllable system EST (PA from A&P 2016b; see also chapter 1), the faithfulness constraints, f.max of f.dep are a class. The extensional trait of onset-requiredness in syllables correlates with P values where the markedness constraint m.Ons is ranked relative to the *subordinate* of the class (6).

6) *EST properties*

P1: m.Ons <> {f.max, f.dep}.sub

α: m.Ons > f.max *OR* m.Ons > f.dep *onsets required*

β. f.max & f.dep > m.Ons *onsetlessness*

The present treatment of classes builds on the significant development in A&P (2016a: §III), and further analyzed in D&P (in prep.). D&P develop the internal structure of classes as uniform branching binary trees. This chapter builds on this, defining classes, κs, as hierarchical tree structures more generally. Hierarchical structure is necessary to recognize internal organization to classes, which are not always unorganized sets. For example, in the PA of nGX (A&P in prep., 2017; also ADP), the P Mult has antagonist {{AFL, AFR}.dom, {Iamb, Troch}.sub}.dom, which includes two classes within a class.

This is an intensional concept of C class, distinct from other concepts of C groupings based on definitional or extensional criteria, such as faithfulness, alignment, or correspondence. To distinguish these, the latter are termed *C families*. Whether the two concepts align depends on the specific system. In PA(EST), the faithfulness Cs f.max and f.dep act as a κ, but markedness Cs, m.Ons and m.NoC, do not.

2.3.1 κ and κ trees[3]

κs are defined as hierarchical structures over subsets of Cs ∈ CON, represented as *κ trees* (7), with a set of *subtrees*, rooted at non-terminal nodes, and leaves labeled with Cs. Non-terminal nodes are labeled with the set of the labels of their immediate child nodes.

7) *Def.* A *κ tree* is a rooted acyclic tree with leaves labeled by Cs ∈ CON and non-terminal nodes labeled by the set of the labels of their child nodes.
A *subtree n* of a κ tree is the tree rooted at a non-root node *n* and all nodes dominated by *n* in κ tree. The edges are those that connect these sub-nodes in κ tree.

A κ is defined as the label of the root node (8). It is the set of the labels of its child nodes, which are themselves subtrees or leaves.

[3] Formal definitions of concepts in this section were developed jointly with Nazarré Merchant.

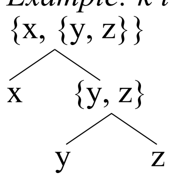
8) *Def.* Given a κ tree, a *class κ* is the label of the root node κ, the set of immediate child nodes of the root. A *sub-κ* is the label of a dominated node *n*, the set of immediate child nodes of the subtree rooted at node *n*.

The immediate child nodes, whose labels form the set naming κ, are the *daughters* (9).

9) *Def.* Given a κ tree, the *daughters* of κ, {d1,…,d*n*}, are the labels of the *n* immediate child nodes of the κ root.

The example κ tree in (10) has root node κ = {x, {y, z}}. The two daughters are d1 = {y, z}, a non-terminal node, and d2 = x, a leaf. The daughters of sub-κ {y, z} are y and z.

10) *Example: κ tree*



A tree has a *height*, the longest chain of edges between the root and a leaf.

11) *Def*[4]: the *height* of tree, *h*, is the number of edges between the root node and the deepest leaf.

A singleton κ, a leaf, has a height of 0. For κ with only terminal daughters *h* = 1; one dominating at least one height-1 sub-κ, *h* = 2, etc. The κ height is always one more than the height of its daughter with the highest *h*. In (10), *h* = 2 for the root-node, dominating one daughter {y,z} with *h* =1 and one *h* = 0.

2.3.2 κ valuations

A κ is the set of daughter node labels in a κ tree, itself a hierarchical structure over a set of Cs ∈ CON. To be interpreted as an antagonist in a P value and converted to an ERC set,

[4] https://en.wikipedia.org/wiki/Tree_(graph_theory).

specific daughters within this set are picked out by operators *dom* and *sub* (12). A κ tree is *valued* when all nodes are assigned an operator, op. A P antagonist is a valued κ, following A&P (2016a: p. 3, (15)).

12) *Def.* A *valuation of κ*, is the assignment of op, dom or sub, to each κ tree node.

13) *Def.* A P *antagonist* is a valued κ, κ.op.

As defined by A&P (p. 2; see also chapter 1), the operators *dom* and *sub* are functions taking a total order, λ, and returning a specific C ∈ CON by virtue of its ranking position relative to the other Cs in λ (14).

14) *Def. dom and sub functions.* For CON, a set of Cs, Ord(CON), the set of all total orders on Cs ∈ CON, the *dom/sub* operator is a function from Ord(CON) → CON where S.dom/sub(λ) = the greatest/least element of S in λ.

A κ is a restrictor on the set of possible outputs to the function, limiting it to the subset of CON that are leaves of a κ tree (15) (following A&P (2016a: (17), (21)). The position of any non-κ-leaf in λ is irrelevant.

15) *Def. κ.op*: Given a valued κ, κ.op, and a linear order λ, κ.op(λ):
   a. If κ is singleton, then κ.op(λ) = C ∈ κ.
   b. If κ is non-singleton, with valued daughters, d1.op1,…,d*n*.op*n*, then, for U = ⋃{d1.op1(λ),…,d*n*.op*n*(λ)}, κ.sub/dom(λ) = lowest/highest ranked C ∈ U in λ.

When κ is a singleton, valuation is trivial: either op returns the same single C. In the representations below, ops are omitted on leaves. For a tree with *n* non-terminal nodes, including the root, there are 2^n possible valuations of non-terminal nodes, a binary choice of op at each of the *n* nodes. A $κ^1$ has two distinct valuations; a $κ^2$ has minimally four (depending on the structure of the daughters). For the κ tree in (10), there are two non-terminal nodes, yielding four possible valuations are those in (16).

16) *Valuations of κ = {x, {y, z}.op1}.op2*

op1 \ *op2*	*dom*	*sub*
dom	{x, {y, z}.dom}.dom x {y, z}.dom y z	{x, {y, z}.dom}.sub x {y, z}.dom y z
sub	{x, {y, z}.sub}.dom x {y, z}.sub y z	{x, {y, z}.sub}.sub x {y, z}.sub y z

For λ = xyz, where linear ordering represents order, the C returned for each of these valuations is shown in (17).

17) *λ = xyz*

a. {x, {y,z}.dom}.dom(λ) = x

b. {x, {y,z}.dom}.sub(λ) = y

c. {x, {y,z}.sub}.dom(λ) = x

d. {x, {y,z}.sub}.sub(λ) = z

There is a set of linear orders that all return the same C, differing in the permutations of both any Cs that are not κ leaves, and of some of the leaves among each other. For example, (d) returns z for any λ in which both x & y > z, regardless of their relative ordering, and that of any other Cs, e.g. wxvyzu, uvyxzw, etc.

In EST, for example, onsets are required under P1.α: m.Ons > {f.dep, f.max}.sub. There are multiple λs for which {f.dep, f.max}.sub returns f.dep, differing in the order of m.Ons, f.max, and m.NoC. The λs consistent with P1.α are all those in which both m.Ons and f.max precede f.dep, with any order between them, and m.NoC.

A&P (2016a:§III) further introduce the distinction of *public* and *private* classes. A class is *public* if its daughters are the antagonists of another P; else it is *private*. This chapter makes a similar but distinct classification of κs as *conflicting* or *non-conflicting*, depending on the conflicting status of their daughters. For a pair of Cs, conflict is defined as the existence of a Border Point Pair (BPP; M&P (70)) in T involving their adjacent transposition (18) (also chapter 3, §3.2.1).

18) *Def: Conflicting Cs:* Two constraints, X and Y, are *conflicting* in T if ∃(Γ1, Γ2) ∈ T, s.t. there is a BPP for Γ1 and Γ2, defined by the adjacent transposition of X and Y: λ1 = P$\underline{XY}$Q ∈ Γ1, λ2 = P$\underline{YX}$Q ∈ Γ2. Else X and Y are *non-conflicting* in T.

The conflicting status of a κ is defined by that of its leaves (19).

19) *Def. Conflicting κs:* A class κ is *conflicting* if for every pair of leaves (X, Y), X and Y are conflicting in T. A class κ is *non-conflicting* if for every pair of leaves (X, Y), X and Y are conflicting in T.[5]

Conflicting status aligns with whether there can or must be a P in the PA antagonizing the daughters. Thus conflicting κs, which generally have such a P, are similar to public classes, and non-conflicting, which do not, to private classes. This relates to *resPs*, the topic of §2.6, which are Ps that antagonize daughters of a κ in another P.

[5] Note: this allows a κ to be neither.

2.3.3 κs and Ps

All Ps antagonize two valued κ.ops, κα.op <> κβ.op (A&P). Values are the ranking conditions resulting from reading the domination relationship in either direction, which generate ERCs.

The operators, dom and sub, have Boolean correlates (A&P 2016a:4, D&P, DelBusso & Merchant 2016, in prep; reviewed in chapter 1). Dom correlates with disjunction when dominant, conjunction when subordinate; sub correlates with conjunction when dominant, disjunction when subordinate. Following from these Boolean relations, multiple P statements can result in logically equivalent P value ERCs, for Ps having distinct κ trees for their antagonists. In (20), the Ps differ in their κβ.ops, but generate the same ranking conditions, shown by converting one to the other using the Boolean distributive law (see also A&P, D&P, DelBusso & Merchant 2016 in prep.).

20) *Logically equivalent P forms*

<table>
<tr><td colspan="2">κβ tree</td><td>{x, {y, z}.sub}.dom
x {y, z}.sub
y z</td><td>{{x,y}.dom {x,z}.dom}.sub
{x,y}.dom {x,z}.dom
x y x z</td></tr>
<tr><td colspan="2">P: w <> κ.op</td><td>w <> {x, yz.sub}.dom</td><td>w <> {xy.dom, xz.dom}.sub</td></tr>
<tr><td rowspan="2">v</td><td>α</td><td>$w > x \wedge (y \vee z)$

LLeW | LeLW</td><td>$w > (x \wedge y) \vee (x \wedge z)$
$= w > x \wedge (y \vee z)$
LLeW | LeLW</td></tr>
<tr><td>β</td><td>$x \vee (y \wedge z) > w$

WWeL,WeWL</td><td>$(x \vee y) \wedge (x \vee z) > w$
$= x \vee (y \wedge z) > w$
WWeL,WeWL</td></tr>
</table>

All Boolean expressions can be converted, using the laws of Boolean algebra, into two normalized forms, disjunctive normal form, DNF, and conjunctive normal form, CNF.

21) *Normal Forms*[6]

a. DNF: the disjunction of conjunctions of literals.

b. CNF: the conjunction of disjunctions of literals.

Similarly, any P antagonist can be converted to a normalized two-level form, in which the outer root node is valued with sub and all daughters with dom, as in the second tree in (20). When subordinated, such a form generates a disjunction of conjunctive ERC sets, similar to DNF (when dominant, it reverses to a conjunction of disjunctions; A&P) [7]. This form is called Dκ.sub (22). It is used in the PvE algorithm to normalize antagonist form for ERC conversion.

22) *Def. Dκ.sub*: a valued κ, κ.op, with a height *h* = 2, where the root is valued with sub and *n* daughter nodes, each valued with dom, {d1.dom,…,d*n*.dom}.sub.

Conversion of any κ.op to a Dκ.sub changes the makeup of the daughters. Conversion uses laws of Boolean algebra, specifically associativity and distributivity, to redistribute and flatten trees (23) (A&P 2016a, in prep., D&P).

23) *Dκ.sub conversion*

a. Associativity:

{{xy}.*sub*, {zw}.*sub*}.*sub* = {{xz}.*sub*, {yw}.*sub*}.*sub* = {xyzw}.*sub*

b. Distributivity:

{x, {yz}.*sub*}.*dom* = {{xy}.*dom*, {xz}.*dom*}.*sub*	*dom over sub*
{x, {yz}.*dom*}.*sub* = {{xy}.*sub*, {xz}.*sub*}.*dom*	*sub over dom*

[6] https://en.wikipedia.org/wiki/Disjunctive_normal_form; https://en.wikipedia.org/wiki/Conjunctive_normal_form.

[7] Conversely, the 2-level form {{}.sub, {}.sub…}.dom is DNF when dominant, CNF when subordinate.

c. κ simplification: Boolean equivalent when dominated

κ.op = {x, {y, {zw}.dom}.sub}.dom	$x \wedge (y \vee (z \wedge w))$	
= {x, {{yz}.sub, {yw}.sub}.dom}.dom	$x \wedge ((y \vee z) \wedge (y \vee w))$	*dist.*
= {x, {yz}.sub, {yw}.sub}.dom	$x \wedge (y \vee z) \wedge (y \vee w)$	*assoc.*
Dκ = {{xy}.dom, {xzw}.dom}.sub	$(x \wedge y) \vee (x \wedge z \wedge w)$	*dist.*

2.4 *Algorithms for generating Γs from PAs*

Ps state ranking conditions antagonizing two valued κ.ops. Each possible value combination, pΓ, is a set of such values. OT ERC Γs are sets of ERCs. This section defines two algorithms for translating between these distinct objects, allowing for their automation.[8] The first, P-values-to-ERCs (PvE), developed by DelBusso & Merchant (in prep.), converts the values of a P to sets of ERCs. The second, Join-Disjunct-Grammars algorithm (JDG), proposed in this dissertation, takes a full pΓ value set and returns a Γ.

2.4.1 Generating value ERCs: PvE

The P-values-to-ERCs (PvE) algorithm takes a P value—antagonized κ.ops—and returns the (set of) ERC sets that characterize it. The algorithm first converts the antagonists to Dκ form (22). It then creates a disjunctive set of ERC sets for each value, consisting of sets of ERCs sharing an L-set that is defined as one of the daughters of the subordinated antagonist. The algorithm is given in pseudo-code form (24), followed by an example of its application.

[8]As of this writing, Merchant has implemented JDG in the PA checker functions in OTWorkplace. PvE is partially implemented.

24) *Pseudo-code of PvE*

Input: P: κα.op <> κβ.op, and CON, a set of Cs that includes all leaves of κα and κβ trees.

0. Convert the antagonists, κα.op and κβ.op, to the logically equivalent forms Dκα.sub and Dκβ.sub, where Dκα.sub has *n* dom-valued daughters, and Dκβ.sub has *m* dom-valued daughters.
1. Generate α ERCs:

For each of the *m* daughters of Dκβ.sub,

Set Eα*i* as an empty set.

For each of the *n* daughters of Dκα.sub, create an ERC, ϵ*j* where

For all Cs in that daughter, add C to W(ϵ*j*)

For all Cs in the Dκβ.sub daughter, add C to L(ϵ*j*)

For all other Cs ∈ CON, add C to e(ϵ*j*)

If NOT(W- and L-sets overlap), then add ϵ*j* to Eα*i*.

End

Return Eα*i*.

End

Set P.α as the disjunction of the *m* Eαs.

2. Generate β ERCs: repeat (1), swapping α/β.
3. Return P.α, P.β.

PvE produces a disjunction of ERC sets for each value. Under P.α, the set consists of *n* disjuncts, one for each daughter of Dκβ.sub, where each disjunct is an ERC set of *m* ERCs, one for each daughter of Dκα.sub. Each of the *n* sets has a distinct L-set, but all share the same *m* W-sets. The L-sets are necessarily distinct, but may be overlapping if the sets of Cs in the daughters overlap. Each of the *m* ERCs in each set has a distinct W-set, but all share the same L-set. Similarly for P.β, switching α/β and *n/m*. If either Dκα.sub or Dκβ.sub is a singleton, then the ERC set produced by PvE consists of a single disjunct where that antagonist is subordinated, as it is the only possible L-set. If both antagonists are singletons, this ERC set contains a single ERC, for the only W-set.

P antagonist classes, κα and κβ, must have non-overlapping sets of leaves. If they overlap, then some C in some ERC must be in both W- and L-sets, which is not a possible ERC, and PvE halts.

PvE is applied in (25). In the input P, κβ.op is a singleton, equivalent to its Dκβ.op form, so P.α generates a single ERC set. For P.β, however, PvE generates a disjunctive set because the dominated antagonist, κα.op, is not a singleton.

25) *PvE applied*

Input: P: {x, {yz}.sub}.dom <> w

0. Convert to Dκv.sub:

   Dκα.sub: {{xy}.dom, {xz}.dom}.sub   2 *daughters*: {xy}.dom, {xz}.dom

   Dκβ.sub: {w.dom}.sub = w   1 *daughter*: w

1. Generate α ERCs:

   For the single Dκβ.sub daughter, w,

   Set Eα1 as an empty ERC set.

   For each of the 2 Dκα.sub daughters, {xy}.dom, {xz}.dom, create an ERC, ϵ*i*, where

   For all Cs in that daughter, add C to W(ϵ*j*): W(ϵ1) = {xy}

$W(\epsilon 2) = \{xz\}$

For all Cs in the Dκβ.sub daughter, add C to $L(\epsilon j)$:

$L(\epsilon 1) = L(\epsilon 2) = \{w\}$

For all other Cs ∈ CON, add C to $e(\epsilon j)$: $e(\epsilon 1) = z$

$e(\epsilon 2) = y$

Add ϵi to Eα1.

End

Return Eα1 = {WWeL, WeWL}

End

Set P.α = Eα1 = {WWeL, WeWL}

2. Generate β ERCs.

For Dκα.sub daughter {xy}.dom,

Set Eβ1 as an empty ERC set.

For each the single Dκβ.sub daughter, w, create an ERC, $\epsilon 1$, where

For all Cs in that daughter, add C to $W(\epsilon 1)$: $W(\epsilon 1) = w$

For all Cs in the Dκβ.sub daughter, add C to $L(\epsilon 1)$: $L(\epsilon 1) = \{xy\}$

For all other Cs ∈ CON, add C to $e(\epsilon 1)$: $e(\epsilon 1) = z$

Add $\epsilon 1$ to Eα1.

End

Return Eβ1 = LLeW.

For Dκα.sub daughter {xz}.dom,

Set Eβ2 as an empty ERC set.

For each the single Dκβ.sub daughter, w, create an ERC, $\epsilon 2$, where

For all Cs in that daughter, add C to W(ϵ2): W(ϵ2) = w

For all Cs in the Dκβ.sub daughter, add C to L(ϵ2): L(ϵ2) = {xz}

For all other Cs ∈ CON, add C to e(ϵ2): e(ϵ2) = y

Add ϵ2 to Eβ2.

End

Return Eβ2 = `LeLW`.

End

Set P.β as the disjunction Eβ1 | Eβ2 = {`LLeW`}|{`LeLW`}

3. Return P.α = {`WWeL`, `WeWL`}, P.β = {`LLeW`}|{`LeLW`}.

For this P, PvE produced one disjunctive value, P.β, with disjuncts differing in L-sets, and one non-disjunctive value, P.α, a conjunctive set of two ERCs with distinct W-sets.

The PvE algorithm is used in the Join Disjunct Grammars algorithm to calculate the full pΓs, combining the sets of ERCs for each of the values.

2.4.2 Join-disjunct-γs (JDG) algorithm

The Join-Disjunct-Grammars (JDG) algorithm produces an ERC Γ from a pΓ. The algorithm uses PvE and fundamental operations of ERC logic, first taking the pΓ set of P values, and converting them to ERC sets. It then calculates the Γ that results, using the Fusional Reduction algorithm (FRed; Brasoveanu & Prince 2011) on the ERC set produced by PvE. This is uncomplicated when the PA does not include any disjunctive Ps. However, if some of the values in pΓ are disjunctive, and do not describe a unique ERC set, they cannot simply be amassed and FRed-ed. JDG draws its name from its treatment of these cases: the algorithm produces a separate ERC set, a γ, for *each* disjunct, then runs FRed on that disjunct in combination with the other P values of pΓ. When there are multiple disjunctive values, a γ is produced for every possible combination of disjuncts from each such value. The algorithm then uses Merchant's (2008, 2011) join operator, ⊕, to join the set of all γs, producing an ERC set.

The join of a set of ERC sets is the smallest ERC set that is separately entailed by each of the joined sets, extracting their shared ranking information (Merchant 2008, 2011). All linear orders, λs, that satisfy any of the sets joined, the *joinards*, also satisfy the join. A join is *conservative* when it is equal to the union of these λs. If, however, there are other λs in T that are consistent with the join but not with any of the joinards, the join is *non-conservative*, being larger than their union. JDG output is not guaranteed to be conservative but conservativity is necessary for PA-hood (§2.5). The algorithm is given in pseudo-code in (26).

26) *Pseudo-code of Join-disjunct-grammars* (JDG)

Input: a set of Ps, {P1,…,P*n*}, and a pΓ, a possible combination of P values, {P1.v1,…,P*n*.v*n*}.

a. For each P value, P.v, of pΓ, generate the P.v ERCs, by PvE.
b. For each of the *m* distinct combinations of disjuncts of each of the P values of pΓ, create an ERC set, γ*i*, consisting of the ERCs of those disjuncts and all value ERCs from non-disjunctive P values in pΓ.
c. Run FRed on each γ*i* ERC set.
d. Join the set of all FRed(γ*i*)s, producing ERC set jΓ.
e. Return jΓ.

When $p\Gamma$ has no disjunctive values, there is a single γ generated by step (b), the sole possible combination. The joining step (d) is trivial in this case, as the join of any Γ with itself is Γ. The example in (27) shows a case with a single disjunctive P, requiring non-trivial use of the join step.[9]

27) *Example: JDG applied*

Input: a set of Ps, {P1, P2}, P1: x <> yz.sub, P2: y <> z, and $p\Gamma$ = {P1.α, P2.α}.

a. For each P value of $p\Gamma$, generate the P.v ERCs (by PvE).

- P1.α is a disjunctive value, with two ERC sets: PvE(P1.α): α1: `WLe` | α2: `WeL`.
- P2.α is non-disjunctive, a single ERC: PvE(P2.α): `eWL`

b. For each of the 2 distinct combinations of disjuncts of P1.α, create an ERC set consisting of that disjunct and the P2.α ERC: γ1 = {`WLe,eWL`}

γ2 = {`WeL,eWL`}

c. Run FRed on each of γ1 and γ2. FRed(γ1) = {`WLL,eWL`}

FRed(γ2) = {`WeL,eWL`}

- γ2 is a subset of γ1. The first ERC in γ1 entails the first in γ2 by L-retraction.

d. Join FRed(γ1) and FRed(γ2), producing ERC set $j\Gamma$.

$j\Gamma$ = {`WLL,eWL`} $\oplus$ {`WeL,eWL`} = {`WeL,eWL`}.

- Since γ2 is a superset, entailed by γ1, it is the smallest ERC set jointly entailed by both; $j\Gamma$ is equal to γ2.

e. Return $j\Gamma$ = {`WeL,eWL`}.

[9] Details of the FRed and join algorithms are not shown here; see Brasovenau & Prince (2011) and Merchant (2008).

While the output of JDG, jΓ, is guaranteed to be an ERC Γ by the join logic, it is not guaranteed to be conservative and non-trivial. It fails to be so in the case in (28) (the invalid cup, §2.6). This example includes the same P1 as in (27), but lacks P2, whose value ERCs resulted in there being a superset γ. Without P2, each γ*i* consists solely of a disjunct ERC set; their join is a trivial ERC, as the L-sets are non-overlapping, and is non-conservative because it includes all Γs in T, larger than the union of the joinards.

28) *Example: non-conservative jΓ*

Input: P1: x <> yz.sub, and pΓ: {P1.α}.

a. Convert values to ERCs. PvE(P1.α): α1: `WLe` | α2: `WeL`.

b. Generate γs: γ1: {`WLe`}

   γ2: {`WeL`}

c. FRed γs. FRed(γ1): {`WLe`}

   FRed(γ2): {`WeL`}

   - Since γs are single ERCs, FRed trivially returns that ERC. Neither is a superset of the other: in γ1, x > y; in γ2, x > z.

d. Join: jΓ = $\gamma 1 \oplus \gamma 2$ = {`Wee`}.

   - The join, jΓ, is trivial.

e. Return jΓ = {`Wee`}.

Conservativity tracks PA validity: if jΓ is non-conservative, the PA is invalid. The ERC set described by that pΓ includes additional λs. This occurs when no disjunct is entailed by, and a superset of, the others. As a result, ranking information from the disjunctive P value is lost in the join; in the case above, the join included no rankings from P1, and so the Γ generated is consistent with either value. The following section establishes that a superset γ ensures conservativity and conservativity means a grammatical partition.

2.5 *Conservativity conditions*

This section presents a set of lemmas establishing the conditions for a jΓ to be conservative. It then proposes a proposition of necessary conditions for a set of Ps to describe a grammatical partition, necessary to be a valid PA that generates an OT T.

A jΓ for a pΓ that includes a disjunctive P value, P1.α, is conservative when it is inconsistent with the opposing value, P1.β (Lemma (31)). Inconsistency is ensured when one of the γs of pΓ, γ*i*, the ERC set using the *i*th disjunct of the set characterizing P1.α (or, with multiple such Ps, the *i*th combination of their disjuncts), is a superset of all other γ (Lemma (32)). Putting these together, a superset γ entails jΓ conservativity (Lemma (33)).

Definitional preliminaries: superset-hood is based λs.

29) *Def. Superset γ*: γ1 is a superset of γ2 iff the set of λs delineated by the γ1 ERC set is a superset of the set delineated by the γ2 ERC set, $\{\lambda |\ \lambda \in \gamma 1\} \supseteq \{\lambda |\ \lambda \in \gamma 2\}$.

Recall that a Γ is an ERC set that defines a set of λs. TOT(ERC set) is a function that returns this λ set for the ERC set argument. For a disjunctive P value, P.v, TOT(PvE(P.v)) denotes the union of the λs sets consistent with any disjunctive set.

30) *Def*: TOT(ERC set) = the set of total orders consistent with the ERC set.

31) *Lemma. jΓ conservativity and inconsistency with P1.β.* Let PA = {P1,… P*n*}, s.t. there is at least one P, P1, where PvE(P1.α) is a disjunctive set of ERC sets,

$\alpha 1|\alpha 2|\ldots|\alpha n$, and let pΓ be a possible value combination, s.t. P1.α ∈ pΓ. Recall that JDG(pΓ) = jΓ. If TOT(jΓ) ∩ TOT(PvE(P1.β)) is empty, then jΓ is conservative.

Proof. The lemma TOT(jΓ1) ∩ TOT(PvE(P1.β)) = {} ⇒ conservative is proven by establishing the contrapositive: if non-conservative, then jΓ is consistent with P1.β (¬conservative ⇒ TOT(jΓ) ∩ TOT(PvE(P1.β)) ≠ {}).

- The values of P1, α & β, partition the set of total orders on CON[10]. P1.α generates the disjunctive set of ERC sets $\alpha 1|\ldots|\alpha n$, so TOT(PvE(P1.α)) is the union of all total orders satisfying any of the ERC sets, α1 to α*n*.
- Suppose P1.α is the only disjunctive value in pΓ. Then the sole locus of variability between the γs joined in jΓ is in the disjunctive ERC sets, α1 to α*n*. All other P value ERCs are shared in all γs and satisfied in TOT(jΓ). If jΓ is non-conservative, then TOT(jΓ) is strictly larger than P1.α and includes a total order, λ, that does not satisfy any of the disjunctive P1.α ERC sets. Because P1 values partition the set of total orders, then λ must be in P1.β. Therefore, TOT(jΓ) ∩ TOT(PvE(P1.β)) ≠ {}, because TOT(jΓ) ∩ TOT(PvE(P1.β)) = λ.
- Suppose there are *m* disjunctive values, {P1.α,…P*m*.α}. The combination of these values is the intersection of the unions of all total orders satisfying any of the disjunctive ERC sets for each P. If jΓ is non-conservative, then TOT(jΓ) is larger than this intersection. Since all other P value ERCs do not differ across γs, jΓ must include a total order, λ, that does not satisfy any of the disjunctive ERC sets of at least one of the *m* disjunctive values, P*i*.α. So λ must satisfy P*i*.β and TOT(jΓ) ∩ TOT(PvE(P*i*.β)) ≠ {}.

[10] Recall that values always partition the entire set of total orders, regardless of scope. Whether Γ has a P value depends on if TOT(Γ) includes total orders consistent with one or both (moot).

Lemma (32) establishes that if there is a superset γ, then γ = jΓ. This follows from the logic of the join, which finds the smallest ERC set that includes all the joinards. A proper superset is not required, so the condition is met if there are multiple equivalent γs that are supersets of the others.

32) *Lemma. Superset γ = jΓ*. Let PA = {P1,… P*n*}, s.t. a subset of *m* P.α values generate disjunctive ERC sets, and let pΓ be a value combination of Ps ∈ PA that includes these disjunctive values. If ∃γ*i*, the γ of pΓ produced by JDG (b) with the *i*th combination of disjunct ERC sets, s.t. γ*i* is a superset of all other pΓ γs, then γ*i* = jΓ.
Proof. From Merchant (2008:101, 2011:12), the join of a set of ERC sets is the smallest set entailed by all. If γ*i* is a superset of all other γs, then it is also entailed by all, and is the smallest such set, as any smaller set would exclude some λ of γ*i*.

The proof that existence of such a γ entails jΓ conservativity follows (Lemma (33)).

33) *Lemma*. *Superset γ ⇒ jΓ conservativity*. If there is a superset γ*i* in the set of pΓ γs, then jΓ is conservative.
Proof. From Lemma (32), if γ*i* is a superset of all other γs of pΓ, then γ*i* = jΓ. Since γ*i* is calculated with a disjunct ERC set from each disjunctive P value in pΓ, the λs satisfying γ*i*, and jΓ, satisfy a disjunct for each disjunctive P value, P.α, and so is inconsistent with P.β: TOT(jΓ) ⋂ TOT(PvE(P.β)) = {} and jΓ is conservative by Lemma (31).

If there is no superset γ*i*, then, since γs differ in the disjunct of some P*i*.α, ranking information from P*i*.α is not retained in the join, jΓ, so inconsistency with P*i*.β—and therefore conservativity—is not guaranteed. An example was given in (28). Building on the above Lemmas, the following Proposition states that if jΓ is conservative for all pΓs, then the set of Ps is a partition of the set of λs into Γs.

34) *Proposition.* Let PA = {P1,…P*n*}, with *m* possible distinct value combinations, {pΓ1,…,pΓ*m*} and CON = the set of Cs that are leaves in the κ tree of any antagonist in a P ∈ PA. If ∀pΓ*i*, the output of JDG(pΓ*i*), jΓ*i*, is conservative, then PA partitions the set of total orders over CON into Γs.

Proof. As joins, all jΓs are ERC sets, so all pΓs generate OT Γs.

To to show that they partition the λ-set, every λ must be in one and only one Γ:

a. The intersection of any two Γs, Γ1 and Γ2 is empty, TOT(Γ1) ∩ TOT(Γ2) = {}.
   Γ1 and Γ2 are produced by JDG(pΓ1) and JDG(pΓ2). As distinct value combinations, pΓ1 and pΓ2 differ in at least one P value, P*x*. By Lemma (31), jΓ for a pΓ with value P*x*.v is conservative if it is inconsistent with P*x*.$\overline{\text{v}}$; by assumption all jΓs are conservative. If pΓ1 includes P*x*.v and pΓ2 includes P*x*.$\overline{\text{v}}$, then every λ in jΓ1 is inconsistent with P*x*.$\overline{\text{v}}$ and every λ in jΓ2 is inconsistent with P*x*.v, so the intersection of their λ sets is empty.
b. All λs are in some Γ, ∀λ, ∃Γ: λ ∈ TOT(Γ).
   Since Ps partition the set of λs, a given total order, λ, is consistent with one and only one value of each P. So there is a set of values, {P1v,…,P*n*v}, s.t. λ is in the ERC set delineated by this set of values. If this value set is instantiated by one of the pΓs, pΓ1, then λ is in jΓ1. Suppose there is no pΓ1 that instantiates the set. Since λ exists, the value set is consistent and cannot be ruled out by contradiction, so it must be eliminated by scope. Then for a pΓ2 that is described by a subset of the values describing λ, λ is in jΓ2, since λ satisfies all these values.

Finally, Lemma (35) establishes that other P values besides the disjunctive value(s) are necessary and must result in a ranking in which κα.op > d*i*.op, for some κβ.op daughter d*i*.op, being entailed in all γs. If there are no other values, then each γ is simply one of the disjunct ERC sets, which are not sub/supersets of each other.

35) *Lemma*. *Other Ps needed*. Let PA = {P1, P2, …, P*n*}, with at least one disjunctive P, P1: κα.op <> κβ.op, so that PvE(P1.α) is a disjunctive set of ERC sets, α1|α2|…|α*n*. If ∃γ*i*, the pΓ γ calculated with disjunct α*i* of P1.α, s.t. γ*i* is a superset of all other pΓ γs, then pΓ must include some value(s) from a subset of {P2,…,P*n*}; they cannot all be moot.

Proof. In the *i*th disjunct ERC set of P1.α, the *i*th κβ.op daughter, d*i*.op, is the L-set. In γ*i*, calculated using this ERC set, κα.op > d*i*.op in all λs satisfying γ*i*.

Since by assumption γ*i* is a superset of all other γs, this ranking must be entailed by all. Each γ differs in P1.α disjunct, which have distinct L-sets, so d*i*.op is *not* the L-set of the ERCs from the P1.α disjunct for any other γ*j*, *j* ≠ *i*. The ranking thus cannot come from P1.α. Therefore, γs must have a value of some other P(s) ∈ PA to establish this ranking.

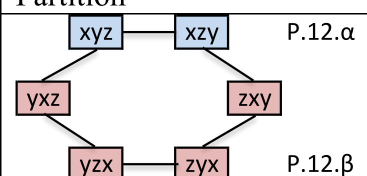
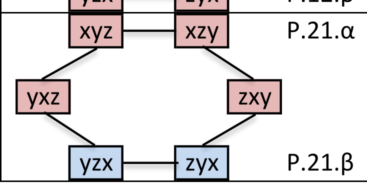
The following section introduces *resPs*, which are used in establishing this ranking.

2.6 *ResPs*

This section introduces *ResPs*, which antagonize the daughters of a κ that occurs in another P. The concept is partly inspired by A&P's (2016a) *public classes*, as public status depends on there being another P in which C in the class are antagonized. A P value ranks some daughter in one antagonist with some daughter in the other, but it does not establish order among the daughters of each antagonist. A resP does so. The term abbreviates 'resolver' because the values assist in 'resolving' a disjunction generated by the P values. For example, the values of P1: x <> {yz}.sub, rank x relative to whichever of y and z is subordinate in a given λ. A resP for {yz}.sub is P2: y <> z, antagonizing the daughters: P2.α: y>z and P2.β: z>y. This is exactly the previously cited case of EST, with x = m.Ons, and y and z = f.max and f.dep.

Before delving into formal details, an example is given in (36). There are two Ps, with different valuations of the same κ. The value ERCs produced by PvE are shown for each, along with their partition of the 3C permutohedron. P.12 values generate single ERCs, with dominator disjunction for β. P.21 values generate multi-ERC sets, conjunctive for β but disjunctive for α, which is satisfied when x dominates *either* of y or z.

36) *P: x <> yz.op*

P	Values	Partition
a. P.12: x <> {yz}.dom	α. WLL β. LWW	
b. P.21: x <> {yz}.sub	α. WLe|WeL β. LWe, LeW	

P.12 is a grammatical partition and a valid PA of the *valid cup* (Merchant & Prince, p.c.). P.12.α is a *top* consisting of two λs, {xyz, xzy}, where x > {yz}, in either order. P.12.β is the complement λ set, {yxz, yzx, zxy, zyx}, all λ in which y or z > x. In neither Γ are y and z consistently ordered in all λs.

In contrast, P.21 does not make a grammatical partition and thus cannot constitute a valid PA. This is the *invalid cup* (Merchant & Prince, p.c.). P.21.β characterizes two λs in which both y and z, in either order, dominate x, {yzx, zyx}, a possible Γ. The complement, {xyz, xzy, yxz, zyx}, cannot be defined by a non-disjunctive ERC set: there is no C that is dominated in all λs. As with the valid cup, no order is established between y and z. This is the case shown in (25) to have a non-conservative JDG output.

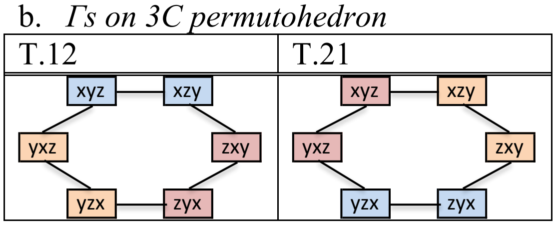
A resP for {yz}.op is P.11 (37). This P antagonizes the two daughters of κ, y and z. Combined (narrow scope) with each of P.12 and P.21 yields the values tables in (38), valid PAs of T.1|2 and T.2|1, respectively. The Γs of each are shown on the 3C permutohedron.

37) *P.11: y <> z*
α. `eWL` β. `eLW`

38) *PA(T.1|2) and PA(T.2|1)*
a. *Value tables*

PA(T.1|2)	P.12	P.11	Γ		PA(T.2|1)	P.21	P.11	Γ
x-top	α		`WLL`		x-bot	β		`LWe, LeW`
y-top	β	α	`LWL`		y-bot	α	β	`WLe, eLW`
z-top	β	β	`LLW`		z-bot	α	α	`WeL, eWL`

b. *Γs on 3C permutohedron*



The Γs of T.1|2 are 'tops', in which a single C dominates the other two, splitting P.12.β of the valid cup; κ daughters, y and z, are ordered in y-top and z-top, as nominally indicated. The T.21 Γs are 'bots', in which a single C is dominated by both the other Cs, splitting P.21.α of the invalid cup; y and z are ordered in y-bot and z-bot, but not in x-bot, where P.21.β generated a non-disjunctive ERC set.

Adding the resP P.11 to P.12 refines a valid PA(T_{OT}), since P.12 alone describes an OT partition. In contrast, P.21 alone does not describe a grammatical partition and requires the resP to be a valid PA(T_{OT}). A κ.dom does not impose the same requirements for validity as a κ.sub. However, the concept of a resP generalizes to both operators, and may be required for a κ.dom to be a valid PA(T_S).

2.6.1 ResPs

ResP-hood is a relation between Ps, as a resP antagonizes daughters of a κ.op antagonist in another P. The definition of a resP is given below. $\hat{P}$ is the set of all daughters of both antagonists: $\{d\alpha 1,\ldots,d\alpha n\} \cup \{d\beta 1,\ldots d\beta m\}$.

39) *Def: resP*: Given a P1: $\kappa 1\alpha.op <> \kappa 1\beta.op$, P2 is a *resP* for a κ1v if:

   a. the daughters of the P2 antagonists, κ2α.op and κ2β.op, are all daughters of some P1 antagonist, $\hat{P}2 \subseteq \hat{P}1$;

   b. $\exists(d1vi, d1vj) \in \kappa 1v$: $d1vi \in \kappa 2v$ & $d1vj \in \kappa 2\overline{v}$.

The definition requires that *all* resP antagonists are daughters of those of P1 ((39)a). For example, for P.21: $x <> \{yz\}$.sub, the antagonists of resP, P.11: $y <> z$ are the set of κβ daughters. A P2: $y <> \{zw\}$.dom is not a resP for P1 because it includes w, not in a P1 antagonist (nor is {zw} as a sub-κ).[11] It further requires that the daughters of the κ1v for which it is a resP be split among the antagonists of P2, and thus antagonized, entailing that P1 ≠ P2 ((39)b).

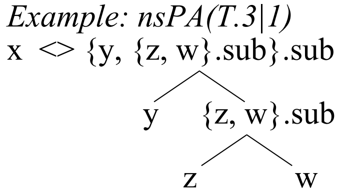
There are multiple distinct P forms that meet the definition. P2 antagonists may be drawn from one or both κ1s, and the subset included in P2 can differ in size, being either *complete*, involving all daughters of a κ1, or *partial*, involving a subset.

[11]See Appendix A for some discussion of this kind of P, a quasi-resP.

40) *Def. Complete resP*: Given a P1 and a resP, P2, for $\kappa 1v \in P1$, P2 is a *complete resP* if $\forall d1v \in \kappa 1v$, $d1v \in \hat{P}2$; else P2 is a *partial* resP.

For example, consider the PA of the 4C bots, T.3|1 (M&P, D&P), which expands T.2|1 by adding a C and a Γ. The structure of the PA is the same as PA(T.21); the relationship that exists between P.21 and P.11 also exists between P.31 and P.21. P.21 is a complete resP for P.31 and similarly for P.11 and P.21. P.21's antagonists are the two daughters of κβ.sub, {y, {zw}.sub}.sub, in P.31. Both P.21 and P.11 antagonize lower nodes in the P.31 κβ tree.

41) *Example: nsPA(T.3|1)*



P.31: x <> {y, {zw}.sub}.sub
 α. `WLee|WeLe|WeeL` β. `LWee,LeWe,LeeW`
P.21: y <> {zw}.sub
 α. `eWLe|eWeL` β. `eLWe,eLeW`
P.11: z <> w
 α. `eeWL` β. `eeLW`

As the value ERCs show, P.31.α generates a disjunctive set, where x dominates one of {yzw}, whichever is subordinate in a λ. P.21 values establish rankings among the daughters of this κ, ordering y relative to the subordinate of {zw}. Finally, P.11 ranks z and w. The value table and ERC Γs are shown in (42).

42) *nsPA(T.3|1) value table*

	P.31	P.21	P.11	ERC Γ
w-bot	α	α	α	`WeeL,eWeL,eeWL`
z-bot	α	α	β	`WeLe,eWLe,eeLW`
y-bot	α	β		`WLee,eLWe,eLeW`
x-bot	β			`LWee,LeWe,LeeW`

The Γ w-bot involves two disjunctive values, with six possible combinations of disjuncts, γs, produced by JDG. Because of the P.11 value, w is in a L-set in all γs, and in jΓ (43). In all γs, x, y, z > w; all except γ6 include additional orderings among {x, y, z}. The first, γ1, defines a total order; γ2 has a 211 structure; γ3, γ4, and γ5 have one additional ranking between two daughters of {x, y, z}. As a result, γ6 delineates a superset of the λs delineated by the other γs, entailed by all and identical to jΓ. The join is conservative.

43) *JDG(ααα) = wbot*

P	γ1	γ2	γ3	γ4	γ5	γ6
31	WLee	WeLe	WeeL	WLee	WeLe	WeeL
21	eWLe	eWLe	eWLe	eWeL	eWeL	eWeL
11	eeWL	eeWL	eeWL	eeWL	eeWL	eeWL
MIB	WLLL eWLL eeWL	WeLL eWLL eeWL	WeeL eWLL eeWL	WLeL eWeL eeWL	WeLL eWeL eeWL	**WeeL** **eWeL** **eeWL**
jΓ	WeeL eWeL eeWL					

The bots are also analyzed with a fully wsPA, using a different form of resP, in which the resP relation is symmetric: two Ps are each a resP for a κv.op in the other. In wsPA(T.2|1), both Ps share the same set of daughters jointly across their antagonists, but distribute them into different antagonists, mapping to distinct κ tree structures. In P1, x is a singleton κ and {yz} a κ^1; in P2, y is a singleton and {xz} a κ^1. While both Ps have a disjunctive value, all consistent combinations produce Γs, using JDG, as shown for αα, the pΓ with two disjunctive values. Of the four γs, one is inconsistent (γ1), while the last, γ4, is the superset, equal to jΓ. A third possible grouping of the three Cs, P3: z <> {xy}.sub, is possible but not necessary to derive the Γs.

44) *Example: wsPA(T.2|1)*

P1: x <> {yz}.sub x <> {y,z}.sub y z	P2: y <> {xz}.sub y <> {x,z}.sub x z
α. WLe | WeL β. LWe, LeW	α. LWe | eWL β. WLe, eLW

45) *JDG(αα) = zbot*

P	γ1	γ2	γ3	γ4
21x	WLe	WeL	WLe	WeL
21y	LWe	LWe	eWL	eWL
MIB	LLe	LWL WeL	WLL eWL	**WeL** **eWL**
jΓ	WeL eWL			

As the example of resPs in (41) shows, more than one resP may be needed when a resP itself generates disjunctive values. A *full resP set, resΦ*, for a κ1v in P1 is a set of resPs whose values jointly order all the daughters of κ1v. This is defined in (46) as a set of resPs such that under each of their possible value combinations, a unique daughter of the normalized form, Dκ1v.sub, is returned as sub.

46) *Def.* Let P1: κ1α.op <> κ1β.op be a P such that PvE(P1.α) generates a disjunctive ERC set. Let {P2,…,P*n*} be a set of resPs for P1. Then {P2,…,P*n*} is a *full resP set, resΦ,* for κ1β if for every allowable, consistent value combination of its values, generating ERC set E*i*, ∃dβ*i*.op ∈ Dκ1β.sub, s.t. ∀λ(E*i*), Dκ1β.sub(λ) = dβ*i*.op.

The definition entails that every pair of daughters must occur in distinct antagonists in some P in the set, as they are ordered under some combination. The same daughter must be returned for every λ consistent with the resΦ value ERC set. For each daughter, there is a value combination defining it as the subordinate. Since all are antagonized in some P in the resΦ, there is a value combination in which each daughter is dominated by all others.

The examples above showed full sets; for nsPA(T.3|1) (42), P.21 and P.11 jointly constitute a resΦ for P.31. Lacking either results in non-conservativity for some pΓ. Example (47)a shows the ERC sets that result from each value combination when P.21 is omitted. The first pΓ, β-, is a Γ (xbot), but jΓs for the others are non-conservative. Under P.11 values, z and w are ranked, but not ordered relative to y. In λs consistent with P.11.α, y and w can occur in any order, including both xzwy and xzyw. But Dκ1β.sub(xzwy) = y and Dκ1β.sub(xzyw) = w, so P.11 is not a resΦ for P.31 because the λs return different Cs. Both pΓs with P.31.α value result in non-conservative jΓs, shown for αα in (47). While γ3 is a superset of γ2, it stands in no such relationship relative to γ1.

47) *No resΦ: invalid PA(T.3|1)*

a. *Value combinations*

P.31: x<>yzw.sub α. `WLee|WeLe|WeeL` β. `LWee,LeWe,LeeW`	P11: z<>w α. `eeWL` β. `eeLW`	jΓ	conservative?
β		`LWee,LeWe,LeeW`	yes
α	β	`eeLW`	no
α	α	`eeWL`	no

b. *JDG(αα)*

P	γ1	γ2	γ3
31	`WLee`	`WeLe`	`WeeL`
11	`eeWL`	`eeWL`	`eeWL`
MIB	`WLee` `eeWL`	`WeLL` `eeWL`	`WeeL` `eeWL`
jΓ		`eeWL`	

2.6.2 ePs

ResPs 'resolve' a disjunction by ordering daughters of a κ; an alternative way of picking out a unique daughter is through entailment. This section examines ePs, which stand in entailment/contradiction relationships with a P1. Entailments occur when P2 antagonists' daughters are drawn from P1 antagonists but P2 does not divide a κ1v across antagonists.

48) *Def: eP*: Given a P1: $\kappa1\alpha$.op <> $\kappa1\beta$.op, P2 is a *eP* for a $\kappa1$v if:

a. the daughters of the antagonists of P2, $\kappa2\alpha$.op <> $\kappa2\beta$.op, consist only of daughters of the P1 antagonists, $\hat{P}2 \subseteq \hat{P}1$;

b. $\nexists$ (d1v*i*, d1v*j*) $\in \kappa1$v: d1v*i* $\in \kappa2$v & d1v*j* $\in \kappa2\overline{\text{v}}$.

The definition parallels that of resPs, sharing the first clause, and negating the second; additionally, both types of Ps are shown below to establish sufficient conditions for conservative JDG outputs, but they do so in distinct ways. To see this, consider the example of ePs in (49). Both Ps share the same antagonist trees, but differ in the valuations of the non-singleton {yz}. This results in entailments between values: P.12.α entails both disjuncts of P.21.α (by L-retraction), and P.21.β entails P.12.β (by W-extension). For a pΓ with P.12.α, either choice of disjunct results in the equivalent output of JDG, and jΓ is conservative, even though the daughters of κ {yz} are not crucially ordered. Note, however, that these two Ps alone are insufficient for a valid PA; the value combination {P.21.α, P.12.β} results in a non-conservative jΓ, as the γs share no Ls.

49) *Example: P.12 and P.21*

P.12: x <> {yz}.dom x <> {y,z}.dom y z	P.21: x <> {yz}.sub x <> {y,z}.sub y z
α. `WLL` β. `LWW`	α. `WLe|WeL` β. `LWe,LeW`

A second example of this type is shown in (50). The Ps share $\kappa\alpha$ (x), but P.11 $\kappa_{11}\beta$ = y is a subset of P.21 $\kappa_{21}\beta$ = {yz}. P.11.α entails the first P.21.α disjunct, and is inconsistent with P.21.β. P.11.β is inconsistent with the first P.21.α disjunct, entailed by P.21.β.

50) *Example: ePs*

P.21: x <> yz.sub	P.11: x <> y
α. `WLe|WeL` β. `LWe,LeW`	α. `WLe` β. `LWe`

As with resPs, there is a concept of a full eP set, based on entailment in this case (51).

51) *Def.* Let P1: κ1α.op <> κ1β.op be a P1 such that PvE(P1.α) generates a disjunctive ERC set. Let {P2,…,Pn} be a set of ePs for P1. Then {P2,…,Pn} is a *full eP set, eΦ* for a κ1v if every allowable, consistent value combination of the set entails or is inconsistent with a disjunct in the PvE(P1.α) ERC set.

2.7 *Sufficient PA conditions for conservativity*

Proposition (34) claimed that complete jΓ conservativity results in a grammatical partition. This section presents a set of new results showing that resPs and ePs establish sufficient conditions for such conservativity. Lemma (35) established that no PA can include only a disjunctive P; other Ps are necessary to rank κα.op > di.op for some di.op for there to be a superset γ. The ranking cannot come from the disjunctive P since each γ is calculated with a different the disjunct.

The needed rankings can be established in two ways: by values that directly rank κα.op > di.op or by values that rank all other daughters of κβ relative to di, so that if κα.op dominates any of them, it transitively dominates di.op. These conditions are met when the PA includes an eΦ or a resΦ, respectively.

Lemma (52) establishes the sufficiency of an eΦ. This case is clear: if the non-disjunctive P values shared by the γs entail one of the disjuncts, then it is also entailed by jΓ, following from the logic of the join. The γ calculated using this disjunct is then entailed by other γs, and is a superset. Conservativity follows from Lemma (33) above. Note that eΦ values may entail multiple disjuncts, so that there are multiple superset γs.

52) *Lemma. eΦ ⇒ superset γ.* Let PA = {P1,…, P*n*}, where P1 is a disjunctive P, so PvE(P1.α) generates a disjunctive set of ERC sets, α1|α2|…|α*m*; and let pΓ be a possible value combination s.t. P1.α ∈ pΓ. If ∃eΦ ∈ PA & pΓ includes eΦ values, then ∃γ*i* of pΓ s.t. γ*i* is a superset of all other γs.

Proof. Recall that an eΦ is the set of *m* Ps, {eP1,…, eP*m*}, such that every allowable, consistent value combination of its values entails or is inconsistent with a disjunct in the PvE(P1.α) ERC set.

Suppose that the pΓ eΦ values generate ERC set E*k* that entails or is solely consistent with P1.α disjunct ERC set α*i*, where d*i*.op ∈ κβ = L-set, so κα.op > d*i*.op. If so, then γ*i*, the γ calculated with α*i*, is a superset.

- *Entailed*: all γs share eΦ values, E*k*. If these entail disjunct α*i*, then all γs entail α*i*. In γ*i*, P1.α ERCs do not contribute any additional rankings not entailed by E*k*. All other γs, calculated with distinct disjuncts, also entail P1.α ERCs in which κα.op dominates some other κβ daughter d*j*.op, so the λs satisfying these γs are subsets of those satisfying γ*i*.
- *Sole consistent*: if α*i* is the only disjunct consistent with E*k*, then γ*i* is a superset of all other γs, because no λs satisfy them, and γ*i* = jΓ.

Lemma (53) establishes sufficiency of a resΦ, through the second means, of ordering the daughters of the κ. The result is that there is some daughter that is dominated by κα.op in *all* γs, and the γ in which it is the *only* daughter so dominated is a superset of the others.

53) *Lemma. res$\Phi \Rightarrow$ superset γ.* Let PA = {P1,…, P*n*}, where P1 is a disjunctive P, so PvE(P1.α) generates a disjunctive set of ERC sets, α1|α2|…|α*m*; and let pΓ be a possible value combination s.t. P1α ∈ pΓ. If ∃resΦ ∈ PA & pΓ includes resΦ values then ∃γ*i* that is a superset of all other pΓ γs.

Proof. Recall that an resΦ is the set of *m* Ps, {rP1,…, rP*m*} s.t. every allowable, consistent value combination of its values, generating ERC set E*i*, ∃d1β*i*.op ∈ Dκ1β.sub, s.t. ∀λ(E*i*), Dκ1β.sub(λ) = d1β*i*.op. Suppose that in pΓ, the resΦ values generate ERC set E*k*, where in all λ satisfying E*k*, κβ.op(λ) = dβ*i*.op. All γs share these values. In γ*i*, the P1.α ERC gives that κα.op > dβ*i*.op.

∀dβ*j*.op, *j* ≠ *i*, dβ*j*.op > dβ*i*.op, since dβ*i*.op is the lowest ranked among the daughters. In each γ*j*, the P1.α ERC generates a ranking κα.op > dβ*j*.op. From E*k*, in all γs, dβ*j*.op > dβ*i*.op, so κα.op also dominates dβ*i*.op by transitivity. Therefore, the λs satisfying each γ*j* are subsets of those satisfying γ*i*, since they also satisfy κα.op > dβ*j*.op.

The Lemmas above are amassed to give the following Theorem on sufficient PA conditions.

54) *Theorem.* Let PA = {P1,…P*n*}, with *m* possible distinct value combinations, {pΓ1,…,pΓ*m*}, and CON = set of all Cs that are leaves of the antagonists of Ps ∈ PA. If, for every P ∈ PA, s.t. PvE(P.v) is a disjunction of ERCs, there is a resΦ or eΦ for P, then PA describes a partition of the set of permutations of CON into OT Γs.

Proof. By Lemmas (52) and (53), the existence of eΦs and resΦs for a disjunctive P in a PA results in superset γs. By Lemma (33), a superset γ results in a conservative jΓ. By Proposition (34), if jΓ is conservative for every pΓ, then the PA describes a partition of the set of permutations of CON into OT Γs.

The Theorem states sufficient conditions for a grammatical partition, but not for an OT partition, which M&P show to be a subset of grammatical partitions. To further guarantee such a partition, the generalized MOAT (GMOAT[12]) of the set of Γs generated must be acyclic, in which case it is an OT partition by M&P, Thesis (177) (p. 110).

2.8 *Summary*

This chapter defined a *valid Property Analysis*, both for a given system, a valid PA(T_S), and as a set of Ps that generates an OT typology, a valid PA(T_{OT}). It then examined the conditions under which a set succeeds or fails to be so.

Assessing PA validity of a set of Ps requires formal methods of translating between the objects of Property Theory—properties and value sets—and those of OT—ERC set grammars and typologies. These conversions have been assumed in previous work, and this chapter sharpens them by developing algorithms to convert values into ERC sets (PvE), and combinations of values into grammars (JDG). The JDG algorithm is of further import because of its role in determining PA validity. When the output is a conservative join for all value combinations, then all of these describe non-overlapping grammars. The set of Ps creates a grammatical partition.

The chapter then developed the concepts of resPs and ePs, formalizing types of relationships between properties in a PA. The proposed Theorem claims that the presence of sets of these Ps in a PA supplies sufficient rankings for the set of Ps to result in conservative JDG outputs and so be a grammatical partition.

---

[12]A generalized MOAT, GMOAT, consists of generalized EPOs (GEPOs) (M&P).

A. *Appendix: quasi-resPs*

While a resΦ is sufficient for JDG conservativity, it is not argued to be strictly necessary because it is sometimes possible to establish the necessary rankings through some combination of rankings from other P values. This appendix examines such a case where the Ps involved are not resPs as defined here; however, they do antagonize daughters of a κ in another P. They fail to be resPs because their antagonists also include additional Cs. These are thus called *quasi-resPs*.

The example in (55) illustrates, using the AOT system QR, which derives from an early iteration of Alber's (2015b) truncation system and Prince's (p.c.) PA thereof, simplifying by removing some Ps to isolate the relations of interest. Here, P.12 is not a resP for P.21 because while its antagonists include those from P.21 κβ, {z, w}, y intrudes, grouped in a κ.dom with z. P.12.β ranks w>z, but P.12.α does not guarantee the reverse, since y is also a possible dominator of w (W-disjunction). However, JDG(ααα) is conservative. The reason is the P.11 value, which establishes x>y. The γ2, using the 2^{nd} disjunct where w is κ.sub, is the superset. Without P.11, the PA is invalid because JDG is non-conservative for αα. Note that in L4, z and w are not consistently ordered in all λs: in 5, z>w, but in the last w>z (xywz). A resP of the form z<>w cannot be used, as it is moot in L4.

55) *PA(T_{GR})*

a. *Ps and value table*

Γ	P.11: x<>y α. `WLee` β. `LWee`	P.21: x<>zw.sub α. `WeLe|WeeL` β. `LeWe,LeeW`	P.12: yz.dom<>w α. `eWWL` β. `eLLW`
L1	β		
L2	α	β	
L3	α	α	β
L4	α	α	α

b. *JDG(ααα) = L4*

P	γ1	γ2
11	WLee	WLee
21	WeLe	WeeL
12	eWWL	eWWL
MIB	WLLL eWWL	**WLeL** **eWWL**
jΓ	**WLeL** **eWWL**	

3 The Structure of Stringency Systems

3.1 *Introduction*

Stringency constraints formalize the notion of markedness scales in linguistic typologies. A *scale* is a linear ordering; markedness scales arise from implicational relations in cross-linguistic comparison. A particular trait *x* is 'more marked' than trait *y* if any language that has *x* also has *y*, but not vice versa.

Much OT work aims to derive the typologies predicted by empirical markedness scales, beginning with Prince & Smolensky's (1993/2004, esp. ch. 8; P&S) analysis of harmonic margin and peak segments in syllables. The markedness of a segment in a position depends on its sonority: more sonorous segments are more 'harmonic' (cross-linguistically preferred) peaks, and less sonorous are less harmonic peaks, v.v. for margins. Generally, if *x* is a possible peak in some language, then so are all more sonorous segments, and reverse for margins. P&S derive this scale through two constraint sets, *P/x and *M/x, violated by segments of different sonority in peak and margin positions, and fix rankings within these sets. *P/t (no stop peaks) universally dominates *P/n (no nasal peaks); the ranking runs in the opposite direction for *M/x constraints.

Subsequent work used the scales to directly generate constraint sets, allowing for full free interaction of constraints in accord with OT logic (see Prince 2002:3-4 on building such Cs by moving from an element to inclusion hierarchy). A reworking of P&S's analysis would have a constraint *P/tn (violated by *both* stop and nasal peaks) in place of *P/n; violation of *P/t entails violation of *P/tn, but not vice versa. *P/tn is said to be *more stringent* than *P/t, because the range of structures violating the former are a superset of those violating the latter.

This chapter develops a formal definition of stringency intrinsic to the structure of OT, defining it as a relation between constraints' *filtration patterns* within a typology, rather than their violation counts or definitions. Constraint (C) definitions are insufficient to determine the existence of stringency, even when intuitively suggesting it, without also knowing the forms they evaluate and the other Cs they interact with. The relation is defined in the context of a full typology. Misalignments can arise between intuitive and formal notions of stringency.

Stringency filtrationally-defined is recognizable by a characteristic MOAT structure, as EPOs in the MOAT represent Cs' filtrations over the grammars (Γs) of the typology. The identification of a MOAT realization of stringency raises the question of how other kinds of constraint relations—conflicting and non—manifest in this structure, the topic of §3.2.1. It further leads to the discovery that the same general EPO signatures can occur in subtler, modified forms in a MOAT. This gives rise to the definition of *partial stringency*: rather than the Cs being stringently related for *all* filtrations, there is *some* filtration product—and thus some Γs—in which the same subset relation between their filtrations is manifested. Partial stringency highlights the complexity of C interactions within a T: multiple kinds of relations can co-occur.

The shared intensional structure of *stringency systems*—those having stringently related Cs in CON—is explicated through Property Analysis. Cs in this relation are *non-conflicting* among themselves, and interact with other Cs in the typology in characteristic ways. The chapter develops the general PA structures of a range of such systems, showing how PAs explicate the relationship. This PA structure recurs in modified form in the cases of partial stringency, along with other PA motifs.

Beyond the inherent interest in understanding the organization of typologies instantiating different types of C relations, the PAs show how intensional typological structure is independent of the linguistic system analyzed, whether phonological, syntactic, or another nature. Extensionally non-equivalent typologies are intensionally equivalent. Intensional structure depends on the filtration patterns of Cs ∈ CON, a particular set of which is realized in stringency systems. Thus when stringency structures are identified in any typology, the same basic units of analysis occur—the properties developed in this chapter.

A typology-based definition connects with the typological nature of scales, which arise from comparison of grammars within a typology. The PAs link a common intensional structure to extensional traits in the languages' optima. The general extensional classification of stringency systems is characterized by a set of inter-related choices regarding the degree to which the scale-defining trait is exhibited in a language's optima: *all*, *none*, or (degrees of) *some*. In the word-order typologies analyzed in the following chapter, languages differ in the degree of complement -head (head-final) order in syntactic phrases in optima: in *all* phrases (all head-final), in *none* (all head-initial), or in *some* phrases, at specific positions in the syntactic structure. This classification recognizes the same set of extensional choices that the Parameter Hierarchies theory argues to structure all syntactic typologies (proposed and developed under the Reconsidering Comparative Syntax project (ReCoS), Roberts 2010, 2012, et seq.). The proposal is compared to PT in the context of the word-order typology in chapter 4.

The chapter is deeply indebted to Prince (2000, 2001), an invaluable source on stringency.

3.1.1 Stringency in OT analyses of linguistic systems

Sets of stringently related Cs, *stringency Cs*, are common in OT analyses of diverse phenomena. The table in (1) surveys several that have been analyzed within PT. The analyses derive scales describing an implicational relationship of the presence of some trait(s) cross-linguistically. The languages of the typologies differ in the degree to which that trait occurs. The table schematizes the empirical markedness scale and lists the corresponding stringency Cs (see the cited sources for full CONs; some C names are modified; current presentation is based on independent examination of the works).

1) *PT Analyzed COT stringency systems*

System	*Empirical implicational scale* *($>_m$: more marked)*	*Stringency Cs* *($>_s$: more stringent)*
S-retraction, simple (Alber 2015a)	sC1 ∈ {k,p,t} $>_m$ sC2 ∈{r,l,n,m,w}	2C: m.C1 $>_s$ m.C1,C2
S-retraction, full (Alber 2015a; 2 interacting scales)	sC1 ∈{k,p.t} $>_m$ sC2 ∈{r,l,n} $>_m$ sC3 ∈{m,w} _s → _ʃ $>_m$ #s → #ʃ	3C: m.C1 $>_s$ m.C1,C2 $>_s$ m.C1,C2,C3 2C: f $>_s$ f.in
Vowel Harmony (D ms. from Tessier & Jesney 2014)	V_{+low} ∈ _σ $>_m$ V_{+low} ∈ #σ (_ = non-initial, # = initial)	2C: m.NoLow $>_s$ m.NoInitLow
Complex stops (Danis 2014; 2 overlapping scales)	c $>_m$ k $>_m$ t kp $>_m$ p $>_m$ t	3C: m.CKP $>_s$ m.CK $>_s$ m.C 3C: m.CKP $>_s$ m.CK $>_s$ m.CKPT
Consonant harmony/ dissimilation (Bennett et al. 2016)	$[d_1t_1]$ $>_m$ $[d_1d_1]$ (subscripts = correspondence)	2C: cc.σEd $>_s$ cc.Id.V
LingPulm Alt (§3.3.3 modifying Bennett 2017)	_N $>_m$ #N (N = nasal click)	2C: f.ons $>_s$ f.init
Agreement with conjuncts (Mitchley 2016, Mitchley & DelBusso in prep.)	Subj: 4h&3h: 3$>_m$ 4 $>_m$ 2; (# = Noun Class) Subj: 8h&8h: 8$>_m$ 2 (8:-h, 2:+h)	3C: CCA $>_s$ dp.NC[1] $>_s$ mx.NC 2C: mx.H $>_s$ dp.I

[1]CCA and dp.NC are in a *partial stringency* relation (§3.2.2), resulting in grammars with anti-Paninian rankings (§3.4.1).

FOFC word order (ch. 4)	TC $>_m$ VT $>_m$ OV	3C: HdL.VTC $>_s$ HdL.TC $>_s$ HdL.C (see ch. for variations)

These systems analyze diverse empirical phenomena, from segmental features to morpho-syntax. However, the results of this chapter show that all systems with stringency Cs have a common intensional structure, elucidated by PAs. The languages are precisely characterized by a set of property values correlating with degree of existence of the marked trait.

This chapter concerns the formal structure of stringency, and much of the following uses AOT systems to distill the core interactions independent of quirks of GEN and CON and to generalize over typologies. Some COT exemplification with Alber (2015a) and variations thereof is used for expository purposes. The following chapter studies a set of specific stringency systems of cross-linguistic word orders, and examines the extensional force of such systems in greater detail.

3.2 *Stringency Constraints*

Stringency is an ordered relationship between constraints[2]. This chapter develops a definition of the relation in terms of *filtration patterns,* which differs from the more common violation sub/superset definition[3]. A filtration-definition is inherently typological. Each C in a typology filters the candidate set, rejecting all candidates to which it assigns non-minimal violations. A C2 is qualitatively more stringent than another C1 if C2 rejectees are a superset of C1 rejectees: C1 allows more candidates

---

[2]The relationship can also exist between a C and a set of other Cs. See Appendix A for cases of this ilk.
[3]Prince's (2000:2) definition: "|G| = |S| + |D|. A constraint G is *more stringent* than S if the violations assessed by G can be partitioned into those assessed by S and those assessed by some other descriptor D = G\S. NB: The violations assessed by S and D must be disjoint."

through its filtration. As filtration patterns are represented in the MOAT, this definition directly connects stringency and EPO structure.

For stringency to exist between C2 and C1, C2 must impose a tighter filtration than C1 over the set of possible optima. In any given hierarchy, C2 and C1 may be preceded by a non-empty set of other Cs whose sequential filtration narrows the set of possible optima. A *decisive* hierarchy determines the optimum (or co-optimum); a *non–decisive* hierarchy does not (2). When evaluating over a UVT, where rows are grammars, filtration by a decisive h results in a single Γ.

2) *Def. Decisive hierarchy*. A hierarchy, h, of Cs ∈ CON is *decisive* iff for K, a cset of violation profile-distinct possible optima, |h[K]| = 1.

In typological or *global stringency*, constraints are in a stringency relationship for all non-decisive hierarchies. A *partial stringency* relation exists when they are stringently-related for only a subset (§3.2.2). The unmodified *stringent* is used throughout for the global relationship. Stringency is first defined relative to a hierarchy (3), and global by universal quantification over all possible hierarchies (4); its properties are examined below.

3) *Def. Filtration stringency.* For a system, S = GEN.S, CON.S;

$\hat{K} = \{K_1,\ldots K_n\}$, the set of all csets of possible optima admitted under GEN.S;

For a pair of Cs, C2, C1 ∈ CON.S, and an h, a (possibly empty) ordered set of Cs ∈ CON.S\{C2, C1}:

a. C2 is *stringent with respect to* C1 relative to h if $\forall K \in \hat{K}$, h.C2[K] ⊆ h.C1[K].

b. C2 is *strictly more stringent* than C1 relative to h if $\exists K \in \hat{K}$: h[K] is non-decisive & h.C2[K] ≠ h.C1[K].

4) *Def. (Global) Stringency*. For a system, S = (GEN.S, CON.S), and a pair of Cs, C2, C1 $\in$ CON.S:

a. C2 is *globally stringent with respect to* C1 in S if $\forall$h, C2 is stringent with respect to C1 ($\forall$h, $\forall$K $\in \hat{K}$, h.C2[K] $\subseteq$ h.C1[K]);

b. C2 is *globally strictly more stringent* than C1 in S if $\forall$h, C2 is strictly more stringent than C1 ($\forall$h, $\exists$K $\in \hat{K}$: h[K] is non-decisive, h.C2[K] $\neq$ h.C1[K]).

Note that when considering a UVT rather than multiple csets, there is single K $\in \hat{K}$. Since the UVT is constructed from the collection of csets, the two objects instantiate the same C relations (Prince 2016a). If C2 and C1 are stringently related in all individual Ks, they are in the UVT; if there is some crucial cset for which they are *not* stringent, then this too is evident in the UVT.

The first clause of the global definition, (4)a, establishes a possibly symmetric relation between Cs. The second, strict stringency (4)b, establishes asymmetric, proper stringency, where for any h, the two C filtrations are not equal in some K$\in\hat{K}$. Where filtrations are equivalent, neither C is more or less stringent than the other—they are equivalent. If this holds for all h and K, then Cs are equivalent in T (see §3.2.1). This chapter focused mainly on the asymmetric, strict relation and uses *stringency* to characterize this. Where two Cs are fully symmetric, they are called *equivalent*; when both relations exist in T, conjunction of terms is used. Equivalence and stringency relations co-exist in complex ways, the full range of which is not analyzed here.

Strict stringency defines a T-internal irreflexive ordering relation over a set of Cs, inheriting the properties of subset relations[4]. For a scale of *n* stringency Cs, {C1, C2,…,

[4]Non-strict relations have subset-equivalence properties: reflexivity, antisymmetry, transitivity.

$C\overline{n}$, Cn} where for each pair ($C\overline{x}$, Cx), Cx is stringent with respect to $C\overline{x}$, Cn is the greatest element, the most stringent, and C1 the least element, the least stringent: $Cn >_s C\overline{n} >_s \ldots >_s C2 >_s C1$.

5) *Stringency ordering relation*

Strict stringency is a binary ordering relation, R, over a set of *n* Cs, {C1,…C*n*}:

a. Irreflexive: $\neg(Cx >_s Cx)$.

- ∀h, K, h.C*x*[K] = h.C*x*[K], failing (3)b; a C cannot be more strictly stringent than itself.

b. Asymmetric: $Cx >_s C\overline{x} \Rightarrow \neg(C\overline{x} >_s Cx)$.

- If $Cx[K] \subset C\overline{x}[K]$, then $C\overline{x}[K] \not\subset Cx[K]$.

c. Transitive: $Cx >_s C\overline{x}$ & $C\overline{x} >_s C\overline{x}\text{-}1 \Rightarrow Cx >_s C\overline{x}\text{-}1$.

- If $Cx[K] \subset C\overline{x}[K]$ and $C\overline{x}[K] \subset C\overline{x}\text{-}1[K]$, then $Cx[K] \subset C\overline{x}\text{-}1[K]$.

The relation is assessed over all filtration products under a hierarchy, h, of $\hat{K}$, the set of all csets of possible optima, K (removing Harmonically Bounded candidates), that arise in the process of optimization. Violation-wise, the set of candidates in h[K] that have a minimal value of C2 is a subset of those having a minimal value of C1, though the minimal values need not be equivalent. The definition derives a survival version of Prince's (2002: 36) 'Satisfaction guaranteed' result: survival of C2 entails survival of C1. A candidate k *survives* C2 if k ∈ C2[K], receiving the minimal number of violations (M&P:78); this need not be 0 if no candidate fully satisfies a C.

6) *'Survival guaranteed': If candidate q survives C2 then it survives C1*:

If $h.C2[K] \subseteq h.C1[K]$, then $\forall k \in K$, $k \in h.C2[K] \Rightarrow k \in h.C1[K]$.

Generalizing, following from the transitivity of the relation, survival of any C*x* entails survival of all less stringent, C*j*, $j < x$, and rejection by C*x* entails such for all more stringent, C*k*, $k > x$. All C2 survivors are thus C1 survivors; C1 cannot distinguish between these, where a C distinguishes two candidates, q1, q2 if C(q1) ≠ C(q2). It only distinguishes among C2 rejectees, so that C1 ranking is only decisive if the set of remaining candidates does not include C2 survivors. That is, where C2 is crucially dominated.

Violation count subset does not entail filtration subset, nor vice versa. For example, McCarthy (2008:65-6) defines stringency as: "Constraint Const1 is more stringent than constraint Const2 if every violation of Const2 is also a violation of Const1, but there are some violations of Const1 that aren't violations of Const2." A main point of breakdown between a filtration and a violation-subset definition arises when C1's minimal value is greater than C2's. In OT, it is how the constraints *distinguish* the candidates that determines the typology, not the exact number of violations assigned.

To see how raw violation counts can obscure relations, consider the violation profiles for the two Cs in (7). C2 assigns a violation to every candidate; its minimal value in this cset is 1. C1's violation counts are always less than or equal to C2's (every violation of C1 is a violation of C2), but C1's *filtration* is a proper subset of C2's (final row).[5] Reducing violations to the minimum value respecting filtration patterns, (1,0,0) for C2 (recomputed VT in (b)) shows the relations. C1 is more stringent than C2 (provided all candidates are possible optima). C1(k) ≤ C2(k) does not guarantee that C2[K] ⊆ C1[K].

---

[5] This situation arises in COT systems where a C cannot be fully satisfied in a cset by GEN. For example, in Danis' (2014) system, the C m.CKPT is violated by all places of articulation, but GEN requires all output segments to have such a place. The system illustrates the distinction between violation count and filtration subsets: on the former, m.CKPT is *more* stringent than m.CKP (violations ≥); on the latter, it is *less* (filtration ⊆).

7) *Violation subset ≠ filtration subset*

a. *Full violations*

K	C1:m.a	C2:m.ab
aa	2	2
a	1	1
b	0	1
C[K]	{b}	{a,b}

b. *Violations, minimized*

C1	C2
2	1
1	0
0	0
{b}	{a,b}

A basic example of a stringency system meeting the definition is shown in (8), a 3C subset inclusion hierarchy, in which the set of structures violating each C*x* is a proper superset of those violating C$\overline{x}$, and a single antagonist, X. This set derives a markedness scale a $>_m$ b $>_m$ c. The lower rows show filtrations for possible non-decisive h[K]s (for h = X: {a}, C3: {d}, C1.X: {b}, C2.X: {c}). X values are the minimal possible for all candidates to be possible optima (if (0, 1, 1, 1), (b) and (c) are harmonically bounded). Lacking X, (d) is the sole possible optimum.

8) *3C inclusion hierarchy stringency scale VT*

K	C1:m.a	C2:m.ab	C3:m.abc	X
a	1	1	1	
b		1	1	1
c			1	2
d				3
h = Ø	{b, c, d}	{c, d}	{d}	{a}
h = C1	--	{c,d}	{d}	{b}
h = C2	{c,d}	--	{d}	{c}

This example underscores the typological contextualization of stringency. In COT, this requires a fully specified system with a complete definition of GEN and Con; in AOT, it requires the complete UVT. Importantly, the relationship must hold not for every possible subset of the possible optima in a cset, but only the filtration products resulting from filtration by a possible h, as these are the sets arising in the course of evaluation and optimization. In the above example, some subsets of K such as {b,c} do not arise under filtration by any h. If it did, the Cs fail the conditions because C2 is more stringent than C3 (filtrations {c}, {b,c}, respectively).

The necessity of considering all csets and filtration products thereof in COT is further brought out by an example. In the Lombardi voicing typology (Lombardi 1999, LVT; also Prince 2000) the C definitions suggest a stringent relation that does not hold under all filtrations. The system contains two faithfulness constraints: a general f.V and a positional f.hd.V, violated, respectively, by change in voicing between input-output correspondents in segments in all or only onset positions (9).

9) *LVT: Constraints*

C	*Def*	*Prose: 1 violation for each (in, out) s.t.:*
f.V	*(i, o): [αV] ∈ i & [¬αV] ∈ o.	in & out have differ in [±V] value.
f.hd.V	*(i, o): [αV] ∈ i & [¬αV] ∈ o & o = σ hd.	in & out have differ in [±V] v & out is a syllable head.

Recall that the definition of stringency must be met for a pair of Cs in all csets in order to obtain in the UVT and in the entire T. In this, f.V and f.hd.V fail. The VT below illustrates the critical cases. In either cset, if the Cs stand at the top of the hierarchy, f.V[K] ⊆ f.hd.V. However, for K /ad.ta/, under h = m.Agr, the subset relation reverses: f.hd.V[K] ⊆ f.V[K]. Furthermore, for K /att.da/, and the same h, the Cs have non-overlapping, conflicting filtrations. Their relative ranking determines the choice between b and c (/att.da/→[att.Ta]~[aDD.da] = eLWL).[6]

[6]Prince identified the omission and its consequence. Full LVT PA: Prince (2013) (also DelBusso 2015 ms).

10) *VT*

Input	Output	m.Agr	f.V	f.hd.V	m.ObV	Comment: unfaithful
ad.ta	a. ad.ta	1			1	none
	b. ad.Da		1	1	2	onset
	c. aT.ta		1			coda
h = m.Agr {b, c}			{b, c}	{c}		
att.da	a. att.da	1			1	none
	b. aDD.da		2		3	2 codas
	c. att.Ta		1	1		onset
h = m.Agr {b, c}			{c}	{b}	{c}	f.hd,V[{b,c}] ⊄ f.V[{b,c}]

As in the crucial csets, so in the UVT, where the effect of m.Agr filtration is clear (11). Over filtration product L3-L7 no subset relationship exists between f.V and f.hd.V (final row).[7]

11) *LVT UVT*[8]

	m.Agr	f.V	f.hd.V	m.ObV
L1	1			2
L2	1	1		1
L3		1	2	1
L4		2		3
L5		1	1	2
L6		3		2
L7		2	1	
h = m.Agr		L3, L5	L4, L6	

These Cs illustrate a case of *partial stringency* (§3.2.2). The definition is met for h = ∅, but crucially *not* for all h.

As the number of possible filtration products grows rapidly as |CON| increases and can be hard to determine even in small systems with complex interactions, deduction from UVT scrutiny becomes infeasible. However, stringency is detectable from the MOAT: EPOs of the Cs in the relation have a characteristic set of properties, the topic of

[7]Violation-subset does not hold in this UVT: in L2, f.hd.V(k) > f.V(k). However, this is fragile: the UVT constructed from a typology calculated with cset /.dad./ added reverses the relation, f.v(L2) = f.hd.V(L2).
[8]Further complications arise in LVT because of a partial stringency (§3.2.2) relation between m.Agr and m.ObV. For present purposes, this is not focused on here.

the next section. The MOAT motif underscores the typological-dependency of the relation, a dependency intrinsic to OT; no relation—conflict, stringency or otherwise—is established in isolation, and, as Grimshaw (p.c.) observes, core relations such as harmonic bounding similarly obtain system-internally, upsettable with change to GEN and/or CON.

3.2.1 Stringency in the MOAT and C (non-)conflict relations

In OT, the main action is in constraint conflict: where two Cs filtrations are non-overlapping, no candidate survives both filtrations, so their ranking determines the optimum. Rankings define Γs. However, in a T, not all constraints conflict with all others. For Cs in a stringency relation, survival of the more stringent entails survival of the less. Non-conflict, however, is not limited to stringency, and other types are examined below. Understanding non-conflicting relations is essential to understanding typological structure more generally.

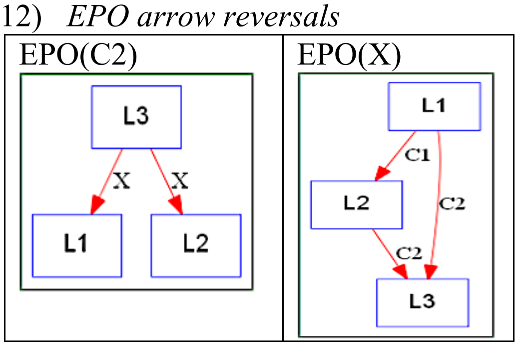
Recall (chapter 1, reviewing M&P) that the MOAT is the collection of EPOs, one for each C ∈ CON, representing its filtration over the set of Γs of T. As M&P show, OT filtration depends on equivalence and order relations only and "the EPO contains the privileged relations that lead to an order on the grammars of the typology" (p. 67). These relations arise from the *border point pairs* (BPPs): two λs belonging to distinct Γs that differ in a single adjacent transposition of Cs (M&P:81 (104)). BPPs are connected to filtrations in that if λ1 = P$\underline{XY}$Q and λ2 = P$\underline{YX}$Q, then for K = UVT, P$\underline{XY}$Q[K] = Γ1 and

$\Gamma 2 \notin P\underline{XY}Q[K]$, and vice versa (i.e., $P\underline{YX}Q[K] = \Gamma 2$)[9]. M&P prove that the MOAT fully determines all Γs in T, so that properties of the MOAT are properties of T itself.

In a given EPO(C), two Γs may be equivalent, $\Gamma 1 \sim_C \Gamma 2$; ordered, $\Gamma 1 \rightarrow_C \Gamma 2$; or non-comparable (unconnected). Equivalence, represented as double blue lines, establishes *equivalence classes* of Γs, Eq_Cs: C does not distinguish among members of the class; they jointly survive or are rejected. Order, shown with a red arrow between Γs, indicates that one, Γ1, survives a filtration from which the other, Γ2, is ejected. Arrows are labeled in the EPO by the C(s) in the BPP(s).

Comparing EPOs shows C relations. If there is an *arrow reversal*, where X and Z order adjacent nodes oppositely, then X and Z conflict because the Cs are in a BPP for those Γs. For example, in the EPOs in (12) (from the system T_{2Core}, §3.3.1.1), there are two such reversals, between L1 and L3, $L3 \rightarrow_{C2} L1$ and $L1 \rightarrow_X L3$, and between L2 and L3, $L3 \rightarrow_{C2} L2$ and $L2 \rightarrow_X L3$. Two constraints Y and W have a *shared arrow* if they both order L1 over adjacent L2, $L1 \rightarrow_Y L2$ and $L1 \rightarrow_W L2$; they evaluate the pair equivalently and either ordering of them in a hierarchy selects L1. Both are in a BPP with an antagonist, for the same Γs. The antagonist's EPO has multiple labeled arrows between L1 and L2, for each.

12) *EPO arrow reversals*



[9]Hereafter, the notation departs from M&P in using h for P for a BPP prefix for consistency throughout and to distinguish from P = property.

C conflict is defined in (13) (see also chapter 2)[10]. As with stringency, conflict is T-dependent. Note that all Cs are ordered in the individual λs ∈ T; whether they *conflict* and are ordered in Γs depends on the distribution of λs into Γs, leading to the presence or absence of Border Point Pairs (BPPs) defined by their adjacent transposition. As with all C relations, conflict is relative to a particular typology T, necessary to determine BPPs. Conflict appears in the MOAT as a direct arrow reversal between adjacent nodes in EPO(X) and EPO(Y).

13) *Def. Conflict*. Two Cs, X & Y, *conflict* in T iff ∃(Γ1, Γ2) ∈ T, s.t. there is a BPP for (Γ1, Γ2) defined by the adjacent transposition of X and Y, λ1= P$\underline{XY}$Q ∈ Γ1, λ2= P$\underline{YX}$Q ∈ Γ2. In the MOAT, $\Gamma 1 \rightarrow_X \Gamma 2$, $\Gamma 2 \rightarrow_Y \Gamma 1$. Else X & Y are *non-conflicting* in T.

If X and Y conflict for every pair of Γs, then they are *global* conflicters in T. Subtleties of this definition arise because the existence of a BPP does not entail that X and Y are ordered the same way in *all* λ(Γ). When the ERC set defining a Γ includes W disjunction, it is satisfied by all λs in which *any*, not necessarily all, Cs in the multi-W-set dominates those in the L-set.

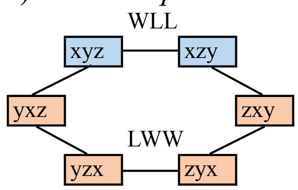
The simple AOT system of the *valid cup*, VC (see also chapter 2) illustrates. T_{VC} contains three constraints, two are which are equivalent, and two Γs mapped to the 3C permutohedron in (14). In the x-top Γ (blue λs) x > y & z (`WLL`); its complement (orange λs) is defined by the ERC `LWW`, where *either* y or z dominates x. Both y and z separately conflict with x by the definition here, with BPPs [xyz, yxz] and [xzy, zxy], respectively.

---

[10]Conflict is defined here in terms of Γs. A. Prince (p.c.) points out that stringency could be defined similarly: for a Γ, a C2 is stringent relative to a C1 if in every h∈Γ, h.C2[K]⊆h.C1[K]. Typological stringency results when this is true for all Γ ∈ T.

But the Γ contains λs in which x > y and in which x > z, so it is not the case that x and y are ordered the same way in all λs.

14) *Valid cup Γs*



Though conflict is a binary relation between two Cs, it does not entail the existence of a binary W/L ERC in Γs in which they conflict for those Cs. BPPs produce *ERCoids* (M&P §3.4), which differ from ERCs in having a fourth value, u for *undetermined*, for Cs in the suffix, Q, of BPP. In fusion, u replaces e as identity, X∘u = X, for X ∈ {W,L,e}. When two ERCoids each have a W where the other has a u, their fusion results in an ERC with W disjunction, i.e., `LWu∘LuW` = `LWW`. Conflict entails the existence of an ERCoid for X and Y, not an ERC.

Cs for which no such BPP exists in T are *non-conflicting*. Stringency is a non-conflicting relationship; as (15) proves, two stringently-related Cs, are never in BPP.

15) *Stringency => non-conflicting.* If C1 and C2 are in a stringency relationship in T, then they are non-conflicting in T.

Proof. Proof by contradiction: assume that C2 and C1 define a BPP.

a. By the def. of stringency, ∀h, K, h.C2[K] ⊆ h.C1[K].

b. By the def. of BPP, ∃(L1, L2) ∈ T: L1 ∈ h.C2.C1.Q[K] & L2 ∈ h.C1.C2.Q[K].

c. Therefore, L2 ∉ h.C2[K] and L1 ∉ h.C1[K], directly contradicting (a), since if h.C2[K] ⊆ h.C1[K], then h.C1[K] cannot contain anything not in h.C2[K].

d. Thus C2 and C1 cannot define a BPP, and by (13) cannot conflict.

Non-conflict is not the only feature of stringency. For a pair of Cs, C2, C1, where C2 is strictly more stringent than C1, C2 $>_S$ C1, EPOs have the four properties in (16).

16) *MOAT features of stringency relations*

a. Equivalence Maintenance: $\forall$pairs of Γs, A, B, $A{\sim_{C2}}B \Rightarrow A{\sim_{C1}}B$.

b. Increased Ordering: $\exists$(A, B): $A{\sim_{C1}}B$ and $A{\rightarrow_{C2}}B$.

c. No conflict: $\forall$(A,B), $A{\rightarrow_{C2}}B \Rightarrow B{\nrightarrow_{C1}}A$.

d. No sharing: $\forall$(A,B), $A{\rightarrow_{C2}}B \Rightarrow A{\nrightarrow_{C1}}B$.

If a pair of EPOs have this set of properties, then the Cs are in a global strict stringency by the filtration definition in (4), as shown in (17).

17) *Deriving stringency from MOAT features*

a. Filtration subset-equality by (16)a and (16)c: C2 survivors survive C1.

i. By (16)a: $\forall\Gamma$ {A,B…} $\in$ h.C2[K]: $A{\sim_{C2}}B$, in h.C1[K], $A{\sim_{C2}}B$. This gives that the survivors of h.C2 are equivalent under h.C1, but does not guarantee that they are the survivors, receiving the minimal value.

ii. Assume the contrary, that they do not survive h.C1[K]. If so, then C2 and C1 must conflict, contradicting (16)c.

- Suppose {A,B} $\in$ h.C2[K] & $\notin$ h.C1[K]. Then $\exists\Gamma \in$ UVT, Z, Z $\in$ h.C1[K] & $\notin$ h.C2[K], since by assumption, equivalent h.C2[K] => equivalent in h.C1.

- Then for any h: {A,B,Z} $\in$ h, h.C2[K] $\cap$ h.C1[K] = $\emptyset$. Minimally, one h meets this condition: h = $\emptyset$.

- If h.C2[K] $\cap$ h.C1[K] = $\emptyset$ then h.<u>C2.C1</u>.Q $\neq$ h.<u>C1.C2</u>.Q, so (C1, C2) are in a BPP, λs of distinct Γs. But by non-conflict (16)c) (C1, C2) cannot be in a

BPP (an arrow reversal), so this cannot hold and so h.C2[K] survivors must also be h.C1[K] survivors.

b. Strictness by (16)b and (16)d: C2 ejects some C1 survivors, ordering them.

   i. From (b) ∃A, B: $A{\sim_{C1}}B$ and $A{\rightarrow_{C2}}B$. This gives that for some h, h.C2[K] ⊂ h.C1[K].

   ii. No sharing (d) ensures that this holds for all h:

   Suppose the contrary, for some pair of Γs (A,B), $A{\rightarrow_{C2}}B$ & $A{\rightarrow_{C1}}B$. Then both C2 and C1 define BPPs with an antagonist, X: h.<u>X.C2</u>.Q[K] = h'.<u>X.C1</u>.Q'[K] = B, & h.<u>C2.X</u>.Q[K] = h'.<u>C1.X</u>.Q'[K] = A.

   If this holds, there is a hierarchy where C2 and C1 filtrations are equivalent.

   - Consider h'.X. First, C2 ∉ h' because both A and B must survive h' if h'.<u>X.C1</u>.Q'[K] = B, and by assumption, C2 rejects B. So C2 ∈ Q'.
   - If h'.X.C2[K] ⊂ h'.X.C1[K], then ∃W ∈ T: W ∉ h'.X.C2[K] & ∈ h'.X.C1[K], so that $B{\sim_{C1}}W$. Since B ∈ h'.X.C1[K], $B{\sim_{X}}W$, giving the following relations for h'[K]:

h'[K]	X	C1	C2	Q'\C2
A	1			…
B		1	1	1(+)[11]
W		1	2	…

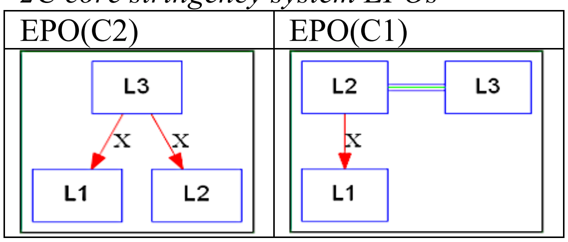
   This mean that h'.C1[K] = h'.C2[K] = A, so for h'[K], C1 and C2 filtrations are equivalent.

This set of properties results in EPO structures in which all Γs in the top Eq_{C2}, the *tops*, are necessarily tops in EPO(C1). These are also unordered, with no arrows in EPO(C1). Order relations in EPO(C1) exist only between non-comparable Γs in EPO(C2). Basic

[11] Since W ∈ T, ∃Y ∈ Q'\C2, $W{\rightarrow_{Y}}B$, else W is harmonically bounded.

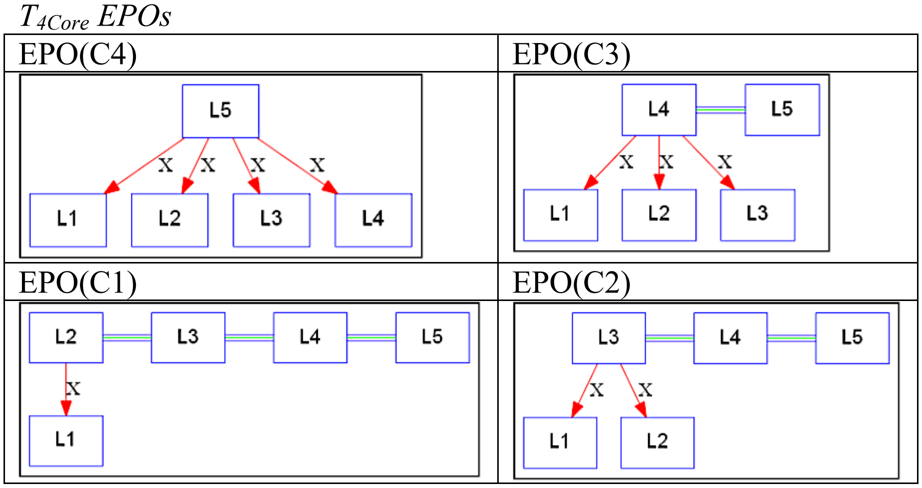
stringency EPO structures are shown in (18), taken from the simplest core AOT system, T_{2Core} (§3.3.1.1). EPO(C2) orders Eq_{C1} = {L2, L3}; EPO(C1) only orders non-comparable Γs in EPO(C2), {L1, L2}. The EPOs represent the filtration subset relationship: each Eq_{C2} is a subset of a corresponding Eq_{C1}. Here, C2[K] = {L3} ⊂ C1[K] = {L2, L3}.

18) *2C core stringency system EPOs*



The transitivity of the stringency relation comes out in the EPOs of larger scales. EPOs of each successive pair display the MOAT correlates, so that for the least stringent, C1, any Γ that is a top in *any* more stringent C EPO is an unordered top in EPO(C1). The EPOs for the system T_{4Core} show this recursive effects with a 4C stringency set.

19) *T_{4Core} EPOs*



While stringency entails non-conflict, the reverse does not hold. Two Cs may be non-conflicting but non-stringency. Cs are *equivalent* if their filtrations are the same: surviving either entails surviving the other (and likewise for rejection). This is the relationship realized by fully symmetric stringency, where the first clause of the definition is met in both directions, for each C. Non-conflicting C filtrations can also lack any cross-entailments, being distinct but partially overlapping (allowing for their joint survival), a relationship here termed *unrelated*. Like stringency, conflict and equivalence have defining MOAT features; the EPOs of unrelated Cs are characterized by the lack of the features of the other kinds of relations, ordering distinct sets.

20) *MOAT motifs of C relations: for a pair of Cs, X,Y:*

a. Arrow reversal: conflict

- If for a pair of Γs (A,B), $A \rightarrow_X B$ & $B \rightarrow_Y A$, then ∃BPP, (h$\underline{XY}$Q, h$\underline{YX}$Q) (by M&P's p. 81 (105): Base relations from a BPP).
- The existence of a BPP establishes conflict by the definition in (13).

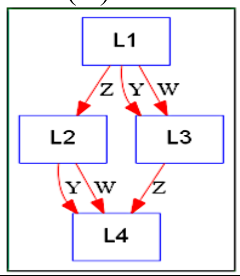
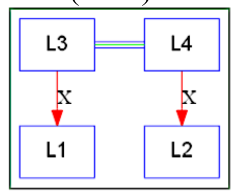
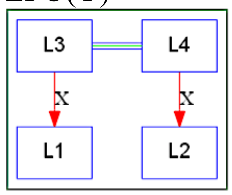
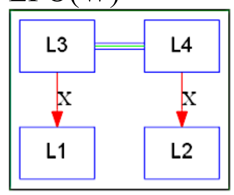
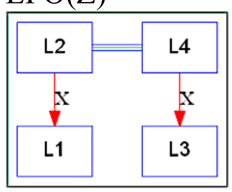
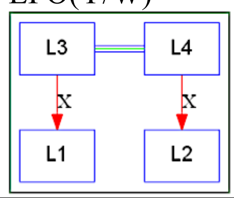
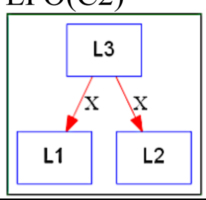
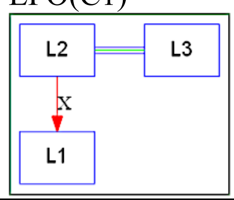
b. Shared order and equivalence: equivalent (identical EPOs)

- If ∀(A,B) ∈ T, if $A \rightarrow_X B \Leftrightarrow A \rightarrow_Y B$ & $A \sim_X B \Leftrightarrow A \sim_Y B$, then all privileged relations are equivalent. Under any h, h.X[K] = h.Y[K].

Each relation has a characteristic PA manifestation, and imposes restrictions on the Ps that must or cannot be in the PA. Conflicting Cs are antagonized in a P, while non-conflicting are not; equivalent Cs are in a κ.dom in all Ps in which they are an antagonist; unrelated Cs occur in separate Ps.

The MOAT and PA structures are shown in (21). EPOs for the first three cases—conflicting, equivalent, unrelated—derive from the AOT system T_{Crel} analyzed immediately below; those of the last—stringency—are repeated from above from T_{2Core} (§3.3.1.1).

21) *Types of C relations*

Type: MOAT structure		*EPOs*		*PA*
Conflicting Arrow reversals: {L1, L3} & {L2,L4}		EPO(X) 	EPO(Y/W) 	Antagonists in some P(s)
Non-conflicting No arrow reversals	*Equivalent* Shared arrows and Eqs	EPO(Y) 	EPO(W) 	κ.dom in Ps
	Unrelated No shared or reversed arrows	EPO(Z) 	EPO(Y/W) 	Separate Ps
	Stringent EPO(C2) orders Eq_{C1}s	EPO(C2) 	EPO(C1) 	κ.op or nsPs

The AOT system generating the X, Y, Z and W EPOs, T_{Crel} (UVT (22)), instantiates multiple C relations. It is analyzed below to illustrate these and their PA manifestations. X conflicts with all other Cs, as the labeled arrows in the EPOs show. Y and W are equivalent in all arrows and equivalence classes. As their filtrations never differ, there is no basis for ranking one versus the other. They conflict, as a class, with X: over all Γs X[K] = L1, and Y/W[K] = {L3, L4}—the arrow reversal between L1 and L3. Under h = Z = {L2, L4}, they also filter differently, resulting in the L2/L4 arrow reversal. X similarly conflicts with Z when h = ∅ or Y/W. Z does not conflict with Y/W: L4 survives filtration by either, as seen in its presence in the top equivalence class of both EPOs. They are not related by conflict, equivalence, or stringency.

22) *T_{Crel} UVT*

	X	Y	W	Z
L1		1	1	1
L2	1	1	1	
L3	1			1
L4	2			

PA(T_{Crel}) has two Ps, antagonizing X and Z, P1, and X and YW.dom, P2. Both are ws; C relations in this system are global, holding under all filtrations and thus all Γs, and neither conflict depends on the other (leading to nsPs). Values combine freely to generate 4 Γs.

23) *PA(T_{Crel})*

P	α	β		P1	P2	MIB
P1: X <> Z	`WeeL`	`LeeW`	L1	α	α	`WLLL`
P2: X <> YW.dom	`WLLe`	`LWWe`	L2	β	α	`LLLW,WLLe`
			L3	α	β	`LWWL,WeeL`
			L4	β	β	`LeeW,LWWe`

The PA has the key features of each relation, listed in the final column of (21). Conflicting Cs are antagonized in a P (X and its two antagonists). Equivalent Cs are a κ.dom in Ps (YW.dom in P2). Unrelated Cs are in separate Ps (Z and YW.dom).

That conflicting Cs must be antagonists in some P(s) aligns with intuition: a P can only generate ERCs in which the conflicting Cs are in opposing L/W sets if they are in distinct antagonists of P. The result cannot be guaranteed from the rankings from other P values in the PA; these may transitively order the Cs by ranking them independently and distinctly with respect to another shared antagonist, but they cannot be ranked directly—as they are in some Γ(s) ∈ T.

The subtleties of the conflict relation discussed above are matched by the definition of P antagonization (24), which requires that the conflicting Cs be in opposite antagonist sets of some P, but not that they be the *sole* members of those sets. They may be in κ.ops with other Cs. Recall that κv is the antagonist set of one side of a P, $\kappa\bar{v}$ that of the other.

24) *Def. P antagonization*. Two Cs, X and Y, are antagonized in P if $X \in \kappa v$, $Y \in \kappa\bar{v}$.

This definition does not guarantee that the P values generate ERCs that non-disjunctively order the C pair. It only entails that for each value, one member of the conflicting pair is in the L-set and one in the W-set for some ERC in the value ERC set, though the W/L-set need not be singleton. This is necessary because P values generate ERCs, not ERCoids. To produce cases of W disjunction, and thus generate Γs as in the valid cup T, T_{VC}, an antagonist must be a κ.dom.[12] In T_{Crel}, X and Y are conflicting ($\underline{XY}[K] \in L1$, $\underline{YX}[K] \in L3$). But if there were a P2': X <> Y, L3 cannot be assigned a value because it also contains the λ WXYZ, where X > Y.

Non-singleton κ.ops are required when Cs act jointly relative to some antagonist. Equivalency is the most basic such relation. When C filtrations are the same, dominance by either results in the same set of surviving candidates, and both must be subordinated for their rejectees to be optimal. These rankings are generated when the equivalent Cs are a κ.dom, as YW.dom exemplifies.

When Cs are unrelated they occur in distinct Ps, though they may have common antagonists.[13] In this case, there is no filtration under which their ranking is decisive, and thus no grounds for their ordering. These kinds of cases arise in COT when Cs assess distinct traits of optima. For example, in the basic syllable system EST (introduced in

---

[12] For L-disjunction, generating a disjunctive ERC set, the Cs must be in a κ.sub. See also ch. 2.

[13] It is sometimes also possible to construct an alternative PA that groups unrelated Cs together if they conflict with the same antagonists. For example, in EST, it is possible to have a P {m.Ons, m.NoCoda}.dom <> {f.max, f.dep}.sub, with the values distinguishing Γs with only .CV. syllables from others. However, the PA must also have either a P antagonizing (non-conflicting) m.Ons and m.NoCoda, or separate Ps antagonizing each of these with {f.max, f.dep}.sub, as in the original PA (A&P 2016b).

Prince & Smolensky, analyzed in A&P 2016b, M&P), m.Ons and m.NoCoda are unrelated in this way: each conflicts with a set of faithfulness Cs to determine the mappings of onsetless and codaful inputs, respectively. In T_{Crel} the unrelatedness of Z and Y/W results in the absence of any Ps antagonizing them.

In the simple system above, the C relations hold over entire T. This is not always the case: two Cs can instantiate multiple relationships within a T. For example, a set of Cs that assess equally under some hierarchy may conflict in another, resulting in the PA features of both relations. The next section examines a case of such complex relations: *partial stringency*, where stringency coexists with another relation.

3.2.2 Partial Stringency

The definition of stringency, and its MOAT detectability, leads to identification of *partial stringency* relationships between Cs. Two Cs are stringently-related for some hierarchy, but not all. The core MOAT motif exists within larger EPO structures.

Such cases are of interest for both their formal properties, and their predictions about the scalar behavior that the stringency systems derive. Prince (2001) analyzes the ways a scale is realized in languages. For a scale $x >_m y >_m z$, a language can distinguish all three levels: z less marked than y, y less marked than x. Conversely, the scale may be fully collapsed in a language, with no distinction made between the categories, $\{x,y,z\}$. Finally, a language may partially collapse a scale, distinguishing some but not all categories. Prince characterizes two different ranking structures that produce distinct collapses. In *Paninian* rankings, where a more stringent C is transitively dominated by a less, the associated languages group together less marked levels, distinguishing only x from $\{y, z\}$. Other coarsenings of the order is *anti-Paninian*, grouping together more marked $\{x, y\}$ in opposition less, *z*, a ranking in which a more stringent C is transitively dominated a less. By the filtration definition of stringency developed here, an anti-Paninian ranking only occurs in the case of partial stringency, specifically, under those hierarchies where the Cs are not stringently related (see §3.4.1 for analysis of an AOT T with anti-Paninian rankings).

Partial stringency is defined below. For partial stringency, quantification over hierarchies is existential, not universal as for global. Global stringency thus entails partial, but not vice versa.

25) *Def. Partial Stringency*. For a system, S = GEN.S, CON.S, and a pair of Cs, C2, C1 ∈ CON.S:

   a. C2 is *partially stringent with respect to* C1 in S if ∃h for which C2 is stringent with respect to C1 ($\exists h: \forall K \in \hat{K}, h.C2[K] \subseteq h.C1[K]$).

   b. C2 is *partially strictly more stringent* than C1 in S if for h, C2 is strictly more stringent that C1 ($\exists K \in \hat{K}$: h[K] is non-decisive and $h.C2[K] \neq h.C1[K]$).

The properties of global stringency relations hold for the h subset where the C filtrations meet the conditions. This corresponds to a subset of Γs ∈ T: those optimal under the relevant filtration(s) of a UVT.

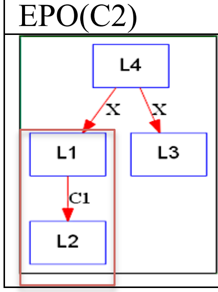
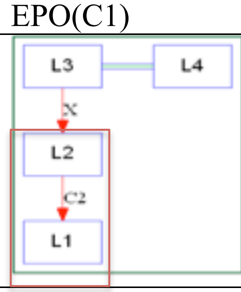
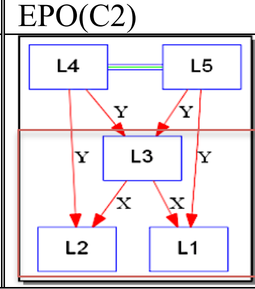
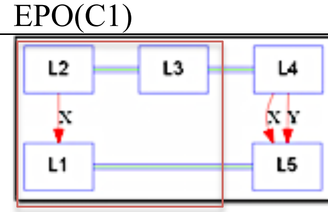
Two cases of partial stringency are classified here: 'lost' and 'derived', relative to the stringency status of a C1 and C2 when h = ∅, unfiltered K. The 'lost' case occurs when the Cs are stringently related for h = ∅, but not for an h' ≠ ∅. These can produce anti-Paninian rankings or cases like LVT, where C1 and C2 conflict under h'. In the reverse case, 'derived' stringency, the relation holds only for a non-empty filtration h', but not h = ∅ (see also Prince & Tesar 2004 for discussion of this kind of relation). If there is no h for which the filtration subset relation holds, no stringency occurs (see §3.2.1 on kinds of relations). The cases are tabulated in (26).

26) *Global and partial stringency*

Type	*Stringent for:*	
	$h = \emptyset$	$h' \neq \emptyset$
global	√	√
partial: lost	√	X
partial: derived	X	√
non-stringent	X	X

In partial cases, the core MOAT features appear embedded or otherwise slightly mangled in the C1 and C2 EPOs. This is shown below, using two cases from AOT systems that are the topics of §3.4.2 and §3.4.3. In the first, (27)a, the stringency structure occurs between L4, L3 and the unordered set of {L1, L2}. However, over {L1, L2} the Cs conflict (arrow reversal), as in LVT above. In the second, (27)b, the EPO structure is embedded, over {L3, L2, L1}. Over the whole T stringency does not hold because C1 orders the equivalence class {L4, L5} of C2. Only among the rejectees of C2, eliminating these Γs, is it more stringent than C1.

27) *Partial stringency system EPOs*

a. *Lost stringency (conflict)*		b. *Derived stringency*	
EPO(C2)	EPO(C1)	EPO(C2)	EPO(C1)
			

Identification and analysis of partial stringency shows how Cs interact in multiple ways within a T. PAs of lost or derived stringency (§3.4) have features of these multiple relations: stringency Ps alongside those characteristic of the other relationship(s).

3.3 *Structure of stringency systems*

Property Analysis (PA) brings out the common core of stringency systems and shows how this C relation plays out in a T. This section develops the structure of basic stringency systems in detail: their Γs and the crucial rankings that define them, explicated by PAs. It then systematically analyzes a set of complications thereof, showing how the same general structures recur.

3.3.1 Core global stringency: T_{nCore}

Since stringency-related Cs are non-conflicting, all Γs are defined by which Cs in the stringency set are ranked relative to which antagonist(s) (along with any other rankings among Cs ∈ Con). The ranking of a C$\overline{x}$ relative to an antagonist, X, depends on the ranking of a more stringent C*x* and X: only in Γs in which X > C*x* are C$\overline{x}$ crucially ranked, since C*x* survival entails C$\overline{x}$ survival. The properties of the PA illustrate this: all stringency set Cs are antagonized with the same antagonists, but those involving a C$\overline{x}$ are dependent on C*x* ranking, either being in a class, κ.dom (chapter 2) or occurring in nsPs.

The most basic stringency system, the AOT system T_{nCore}, has a single antagonist, X, and a set of *n* stringently-related Cs. This system is directly instantiated in the COT systems of Alber (2015a; simple) and FOFC (chapter 4), presumably among many others in the literature. Its structure lies at the core of all stringency systems.

3.3.1.1 T_{2Core}

The simplest T_{nCore} system is T_{2Core}, with two stringently-related Cs, C2, C1. This system is first shown through a COT instantiation—Alber's (2015a) simple system—then an AOT system is used to generalize and further explicate the structure.

Alber's (2015a) simple system derives the scalar generalization that the degree of s-retraction in modern Germanic languages, s → ʃ before a consonant, depends on the consonant's sonority: if s-retraction (SR) occurs before a sonorant (sn → ʃn), then it occur before an obstruent (sk → ʃk) but not vice versa.

The basic SR system contains two stringently-related markedness constraints, defined in (28)b (from Alber 2015a:7). These are violated by s-consonant /sc/ clusters, but not by /ʃc/ clusters, based on the sonority of the following consonant. Where an input contains the violating structure, unfaithful mapping of s, retracting to ʃ satisfies the markedness Cs. More stringent m2 is violated by any such cluster; less stringent m1 only when the following consonant is an obstruent. Their antagonist is a general faithfulness constraint, violated by unfaithful segmental mappings. For each /sc/ input, GEN (28)a produces both faithful and retracted s candidates (all else is faithful).

28) *SR GEN and Con*

a. *Gen*: Inputs[14]: /sc/, c ∈ {k, n}, k = [-sonorant], n = [+sonorant]

Outputs: {sc, ʃc}, $c_{out} = c_{in}$.

b. *Con*: m2: *{sk,sn} (m.kn)

m1: *sk (m.k)

f: *(S_{in}, S_{out}): S_{in} = s & S_{out} = ʃ.

The two inputs in the VT (29)a, a Universal Support, produce the UVT in (29)b, annotated with the degree of SR in the language: all, some (before obstruents only), or none. These C meet the definition of stringency: f-filtration is decisive; when h = ∅, m2 filtration, {L3}, is a subset of m1 filtration, {L2, L3}.

[14]Inputs of the form /ʃc/ are trivial: the faithful mapping candidate is the single possible optimum.

29) T_{SR}

a. *VT*

Input	Output	f	m2.kn	m1.k
sk	sk		1	1
	ʃk	1		
sn	sn		1	
	ʃn	1		

b. *UVT*

	f	m2.kn	m1.k	*Γ*	*SR*
L1		2	1	f > m2, m1	*none (most 'marked')*
L2	1	1		m1 > f > m2	*some: _k, not _n*
L3	2			m2 > f	*all (least 'marked')*

The minimal UVT (mUVT, chapter 1 (9)) reduces m2 violations to 1 for both Γs. The general mUVT for any T_{2Core} system is given below, along with the Γs it generates. ERCs are given in VT order, with '.' separating X from C2, C2.

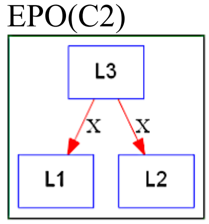
30) *T_{2Core} UVT, MOAT and Γs*

a. *UVT*

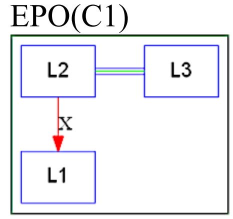
	X	C2	C1
L1		1	1
L2	1	1	
L3	2		

b. *MOAT*

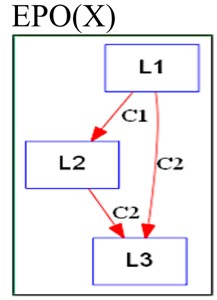
EPO(C2)



EPO(C1)

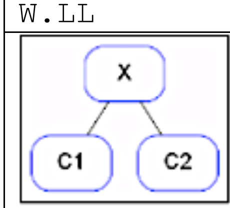


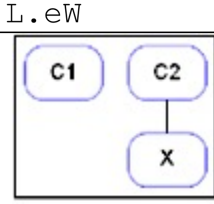
EPO(X)



c. *Γs (ERC order: X.C2.C1)*

L1 (2λ)	L2 (1λ)	L3 (3λ)
`W.LL`	`L.WL,W.eL`	`L.eW`

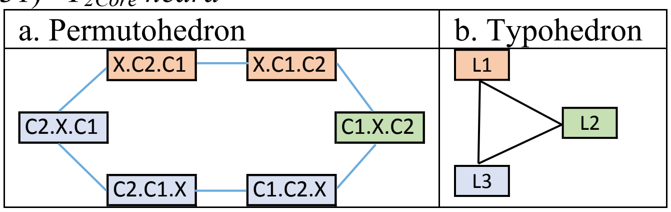




L3 encompasses 3 λs, half of the total orders of the permutations of CON. Only C2 and X are ranked. L1 consists of 2 λs, allowing either ordering of C1 and C2 (both dominated).

L2 is a single total order (1 λ). The languages are mapped to the 3C permutohedron below; combining the nodes of each Γ yields a triangular typohedron.

31) *T_{2Core} hedra*



In all Γs, C2 and X are crucially ordered, and their ranking fully defines L3. In T_{SR} this is the language with *all* SR, before all consonants, the *least* marked (see §3.5 and chapter 4 on extensional traits of stringency system languages, and PA classifications). In PA(T_{2Core}), the ranking is generated by the values of P2, splitting T_{2Core} as in the value table (32). L1 and L2 share P2.α; their T_{SR} correlates share having some degree of faithfulness (*no* SR for some inputs).

32) *PA(T_{2Core})*

a. *Properties*

Property		*Value ERCs*	
		α	*β*
P2	X <> C2	`WLe`	`LWe`

b. *Value table 1*

	P2
L1	α: `WLe`
L2	α: `WLe`
L3	β: `LWe`

C1 and X are only crucially ranked in those Γs where C2 > X (L1, L2, the P2.α Γs) (33).

33) *T_{2Core} Rankings*

- P2.β: C2 > X *no C1 & X ranking (L3)*

  Since ∀q∈K, q∈C2[K] => q∈C1[K] (C2 survival => C1 survival), C1 does not distinguish among q∈C2[K]: all receive minimal value; C1 ranking has no effect.

- P2.α: X > C2 *C1 & X ranking (L1, L2)*

C2 is crucially dominated so C2[K] ∩ X[K] = ∅. X and C1 conflict in Γs ∉ C2[K] (rejectees), so their ranking is decisive.

- C1 > X (> C2): C1 transitively dominates C2 (Paninian ranking) (L2).
- X > C1: X dominates both C1 and C2 (no crucial ranking between them) (L1).

Because C1 and X are not ranked in all Γs, they cannot be sole antagonists in a wsP. In wsPA(T_{2Core}), C1 is in a κ.dom with C2 in P1. C1 and X ranking thus directly involves C2 ranking. Since C2 and X are also ranked in P2, the values of the two Ps entail/contradict each other. As a result, C1 is unranked in L3, and the 4th logical value combination is a contradiction: the value ERCs are inconsistent and fuse to L+ (Prince 2002). The full wsPA is in (34).

34) *ws PA(T_{2Core})*

a. *Properties*

Property		*Value ERCs* α	*Value ERCs* β
P2	X <> C2	WLe	LWe
wsP1	X <> {C2,C1}.dom	WLL	LWW

b. *Value table*

	P2	wsP1	Ranking
L1	α: WLe	α: WLL	X>C2 & C1
L2	α: WLe	β: LWW	C1>X>C2
∅	β: LWe	α: WLL	X>C2 & C1, C2>X
L3	β: LWe	β: LWW	C2 > X

The P value ERCs values derive the Γs through their entailments and contradictions:

- P1.α => P2.α, by L-retraction: if X > C2 & C1 (P1.α: WL**L**), then X > C2 (P2.α: WL**e**) (L3).
- P2.α is consistent P1.β: C1 and C2 are ranked differently wrt X (MIB: LLW, WLe) (L2).

- P2.β => P1.β, by W-extension: W(P2.β) = C2 is a singleton, requiring C2 > X (LW**e**) while the dominant κ.dom in P1.β generates an ERC with multiple W's (C2 *or* C1 > X; LW**W**) (L1).
- P2.β+P1.α fuse L+ (LWe∘WLL=LLL), since P1.α => P2.α and P2.β => P1.β. These values require both C2 > X and X > C2.

The same result of C1 and X non-ranking in L1 is achievable by limiting the *scope* (chapter 1) of the property antagonizing them, defining the space of their conflict by the ranking X > C2 (P2.α). The scope of nsP1, Σ(nsP1) is {L1, L2}, excluding L3, where it is moot: each ordering of its antagonists occurs in some λ(L3). In nsPA (35), possible value combinations are restricted by mootness, not contradiction. This PA structure uses the filtration entailments between the Cs themselves. The treeoid shows the nested choices: C1 and X ranking occurs only under P2α.

35) *nsPA(T_{2Core})*

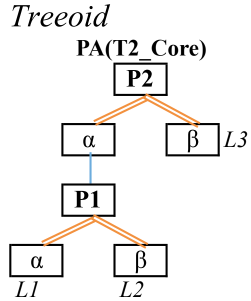
a. *Properties*

Property		*Value ERCs* α	*Value ERCs* β	*Scope*
P2	X <> C2	WLe	LWe	
nsP1	X <> C1	WeL	LeW	/P2.α

b. *Value table*

	P2	P1
L1	α: WLe	α: WeL
L2	α: WLe	β: LeW
L3	β: LWe	

c. *Treeoid*



3.3.1.2 Scaling up: T_{nCore}

The results of the previous section generalize systems with larger sets of stringent Cs. As defined above, a T_{nCore} is a system with *n* Cs in the stringency set, and a single antagonist X. T_{nCore} has *n*+1 Γs, each a 2- or 3-level ranking structure, with all crucial rankings between X and members of the stringency set. In any Γ, if C*x* > X, then all more stringent Cs, C*i*, *i*>*x*, are dominated by X, and all less stringent, C*k, k*<*x*, are not crucially ranked.

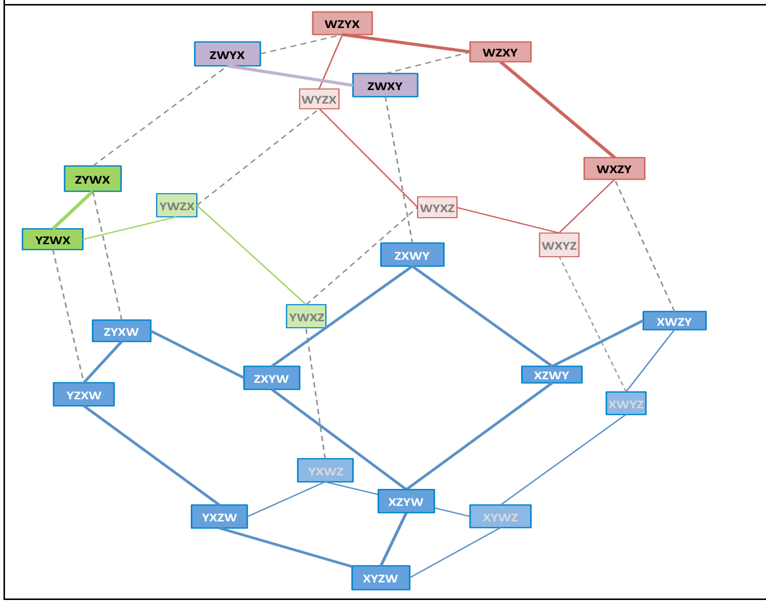
There are two 2-level Γs. L*n* is defined by C*n* > X; C*n* and thus all other stringency Cs are satisfied; all C*k*, *k*<*n* are freely ranked in the Γ. In L1, X dominates all stringency Cs. The languages of these Γs realize the extremes of the scale: *none* and *all* options, resp., for the occurrence of the relevant marked trait. Other Γs generate languages that have some degree of markedness in optima. Each L*x*, for 1<*x*<*n*, is a 3-level ranking structures, with stringency Cs ranked on either side of X, where C*x* > X and all C*i, i*>*x* are dominated.

- L*n* covers half of the λs ((*n*+1)!/2) of the permutohedron;
- L1 covers *n*! λs, an *n*-dimensional shape on the permutohedron;
- Other Γs have fewer λs, 3 levels of ordering.

The typohedron for a T_{nCore} is an *n*+1 (=|Con|)-dimensional object in which all Γs are adjacent (an *n-simplex,* h/t A. Prince). The permutohedron and typohedron of T_{3Core} are shown below.

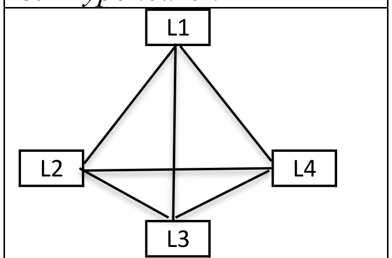
36) *T_{3Core} Hedra*

a. *Permutohedron*



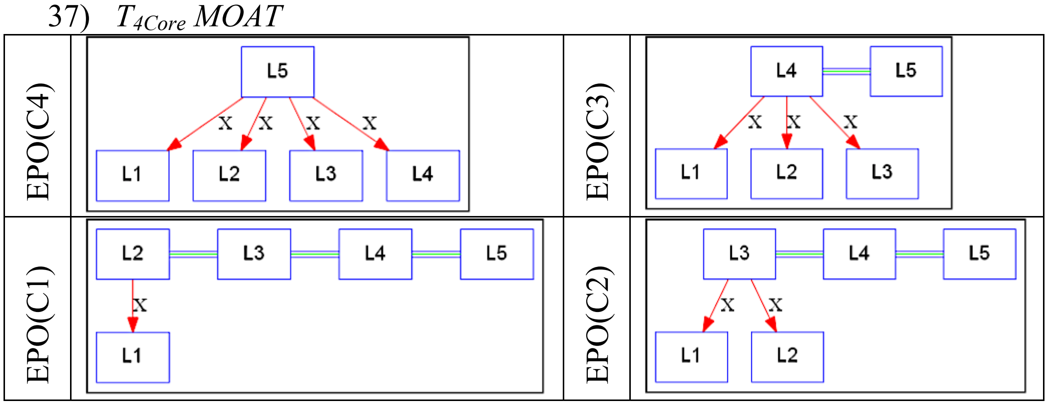
Constraints:	*Γ*:
x = C3	L1 = red
y = C2	L2 = purple
z = C1	L3 = green
w = X	L4 = blue

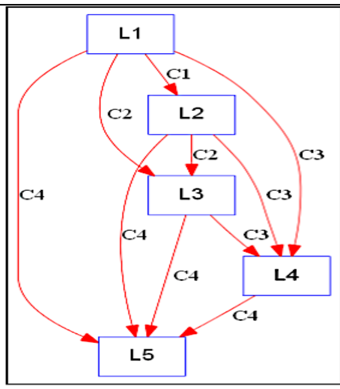
b. *Typohedron*



The MOAT structure repeats successively across the EPOs: for each pair in the stringency set, EPO(Cx) orders some Γs in an $Eq_{C\bar{x}}$; any top Γs in EPO(Cx) 'float' in EPO($C\bar{x}$), as shown in the T_{4Core} MOAT below. EPO(X) shows that every Γ has some arrow labeled with C4—all require ranking of X and the most stringent C—while only {L1, L2} have a C1-labeled arrow—only in these is the least stringent C crucially ranked.

37) *T_{4Core} MOAT*



EPO(X)	

PA(T_{2Core}) structures generalize to PA(T_{nCore}). In wsPA(T_{nCore}), each P antagonizes X with a κ.dom, where κ is a subset of the *n* stringency Cs. All κs include C*n*, the greatest element wrt the stringency ordering; it is the sole member in P*n*. In all P*x*, C*x* defines the *lower* bound of the set in κ: all Cs more stringent than C*x* are in κ. For example, in P3, κ = {C*n*,…C3}. The *less* stringent the C*x*, the *larger* the κ set. The PA is schematized in (38); there are *n* Ps, one for each member of the stringency set, where it is the lower bound of the κ set (ERC order: X.C*n*…..C1).

38) *ws PA(T_{nCore})*

a. *Properties*

∀C*x*, ∃P*x* ∈ wsPA(T_{nCore}): X <> κ.dom, κ = {C*n*,…,C*x*}.

Property		*Value ERCs*	
		α	*β*
P*n*	X <> C*n*	WLe…ee	LWe…ee
P$\overline{n}$	X <> {C*n*, C$\overline{n}$}.dom	WLL…ee	LWW…ee
…	…	…	…
P2	X <> {C*n*, C$\overline{n}$, … C2}.dom	WLL…Le	LWW…We
P1	X <> {C*n*, C$\overline{n}$, … C2, C1}.dom	WLL…LL	LWW…WW

b. *Value table schematized*

	P*n*	P$\overline{n}$	…	P1	Γ	Ranking
L1	α	α	…	α	WLL…LL	X > C*n*…C1
L2	α	α	…	β	WLL…Le, LLL…LW	C1 > X > C*n*…C2
…	α	α	…	β	WLL…ee, LLL…We	C$\overline{x}$ > X > C*n*…C*x*
L$\overline{n}$	α	β	…	β	WLe…ee, LLW…ee	C$\overline{n}$ > X > C*n*
L*n*	β	β	…	β	LWe…ee	C*n* > X

The P value ERC entailments and contradictions from wsPA(T_{2Core}) hold for consecutive pairs of Ps, following from the sub/superset relation of the antagonist sets of the Ps. A dominant κ.dom generates an ERC with dominator disjunction, a W for each C∈κ; a subordinate κ.dom generates an ERC with subordinate conjunction, an L for each C∈κ (see chapter 2). As each κ_x.dom ∈ Px is a subset of $\kappa_{\overline{x}}$.dom ∈ P$\overline{x}$, there is a sub/superset relation between L- and W-sets of the value ERCs. Each Px.α => P$\overline{x}$.α, by L-retraction, and each P$\overline{x}$.β =>Px.β, byW-extension (39).

39) *Entailments between P values*

- P$\overline{x}$.β => Px.β — LW**e**…ee => LW**W**…ee — *W-extension*
- Px.α => P$\overline{x}$.α — WL**L**…ee => WL**e**…ee — *L-retraction*
- P$\overline{x}$.β ∘ Px.α = L^+ — **LW**e…ee ∘ **WL**L…ee = LLL…ee — *fusion = L^+*

Contractions and entailments eliminate many of the 2^n logically possible combinations of wsP values. Only n+1 generate Γs; all others result in ranking contradictions, with a subset of the value ERCs fusing to L^+. Each Px only splits one value of P$\overline{x}$ (P$\overline{x}$.β) (conversely, P$\overline{x}$ only splits Px.α). The possible value combinations are shown in the value table above. When listed from Pn to P1, all Γs are defined by a sequence of 0 or more α values followed by 0 or more β (α*β*); once a Γ has Px.β, it has Pi.β for all Pi, $1 \leq i < x$.

Generalizing from nsPA(T_{2Core}), nsPA(T_{nCore}) (40) antagonizes each Cx of the stringency set with X in a separate Px, whose scope is defined by a value of Px+1. As in wsPA, there are n Ps, one for each C in the stringency set. The scopal structure is a uniform-branching treeoid, aligning with the scale: extensionally, the languages defined at the top and bottom of the treeoid realize the extreme options, having *all* or *none* of the marked trait, respectively (§3.5 and chapter 4). The treeoid extends the T_{2Core} treeoid, iterating the same structure over a larger set of Ps.

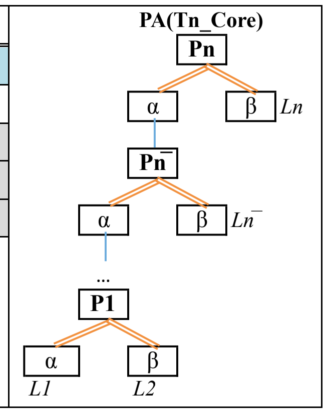
40) *nsPA(T_{nCore})*

a. *Properties:* ∀C*x*, ∃P*x* ∈ nsPA(T_{nCore}): X <> C*x*, Σ(P*x*) = P*x*+1.α.

Property		*Scope*	*α*	*β*
P*n*	X <> C*n*		WLe…ee	LWe…ee
P$\overline{n}$	X <> C$\overline{n}$	/P*n*.α	WeL…ee	LeW…ee
…	…	…	…	…
P2	X <> C2	/P3.α	Wee…Le	Lee…We
P1	X <> C1	/P2.α	Wee…eL	Lee…eW

b. *Value table and treeoid schematized*

	P*n*	P$\overline{n}$	…	P1
L1	α	α	…	α
L2	α	α	…	β
…	α	α	…	
L$\overline{n}$	α	β	…	
L*n*	β		…	



In eliminating κ.dom, nsPA eliminates dominator disjunctions and subordinate conjunction and thus the value entailments; possible combinations are limited to exactly *n*+1 by scope (mootness), not contradiction. A Γ has a P*x* value iff P$\overline{x}$.α ∈ Γ. Γ are defined as a sequence of 0 or more α values followed by 0 or 1 β, α*(β); if a Γ has P*x*.β, it is moot for all P*i*, 1≤*i*≤*x*.

The two versions of PA(T_{nCore}) are compared in the table below. They generate the same Γs using different antagonist sets and scopes. In both, each P involving a C*x* is replicated for C$\overline{x}$ (ws or ns Ps). The same structure persists when X is a non-singleton κ.op, or if C*x* is in a κ.op with some other non-stringency set C; C$\overline{x}$ Ps replicate these (Bennett & DelBusso in prep. analyze such a case).

41) *PA(T_{nCore})*

	κ	*Scopes*	*Value combination limitation*
wsPA	∀Cx, ∃κ: {Cn, …,Cx} ∈ κ	wide	contradiction: …βα… = L^+
nsPA	(all singleton)	Σ(Px) = Px+1.α	mootness

3.3.2 Complexifying

T_{nCore} is the simplest system instantiating stringency relations. The typological structure changes in systematic ways when either the antagonist set or the scale itself is expanded. This section analyzes several expansions, showing how each manifests in different modifications of the core PA. The systems examined change T_{nCore} by either altering the antagonist(s) ($T_{nCoreXY}$, $T_{n\times m}$) or the scale Cs, with two overlapping stringency sets (T_{nmCore}). Each system makes a single change to fully understand its implications, though they may of course coexist. The AOT systems analyzed have COT instantiations, noted in the last table column below.

42) *Variations on T_{nCore} stringency systems*

AOT name	*Description*		*COT examples*
	Scale	*Antagonist*	
T_{nCore}	1, *n*Cs	1 (X)	Alber (2015a), simple (n=2); ch. 4 FOFC (n=3)
$T_{nCoreXY}$		2 (X, Y)	LingPulmAlt, (n = 2) (§3.3.3)
$T_{n\times m}$		Scale, *m*Cs	Alber (2015a), complex (n=3, m=2)
T_{nmCore}	2, *n*C, *m*C, overlapping	1 (X)	Danis (2014) (n/m = 3, 2C overlap in more stringent).

3.3.2.1 Multiple antagonists: PA(T_{nCoreX}) × PA(T_{nCoreY})

When the stringency scale Cs conflict independently with multiple antagonists, the core PA replicates for each. Each antagonist that interacts with the stringency Cs defines a *subPA* with a PA(T_{nCore}) structure. A *subPA* is a subset of Ps in a PA, involving the interaction of a subset of Cs, that acts independently of the other Ps in the PA. Any nsP is in a subPA with those Ps that define its scope. With COT systems, the subPA value combinations often determine the optimal mappings for a specific set of inputs (see also Bennett & DelBusso to appear, using 'subsystem'). The full PA is the union of the subPAs; the set of Γs generated is the product of their consistent value combinations.

The AOT system $T_{2CoreXY}$ is constructed to realize the multiple-antagonist structure, with X and Y each interacting with C2 and C1, but not with each other. In the mUVT (43)a, X and Y each establish three 'blocks' of grammars (distinguished by bolded lines for X), receiving 0, 1 or 2 violations, parallel to the three-way divide of X in T_{2Core}. The EPOs (43)b show the stringency characteristics: each equivalence class in EPO(C1) is ordered in EPO(C2). The same structures that hold between the three Γs of T_{2Core} recur between {L9, L8, L7} and {L3, L2, L1} with Y, and {L9, L6, L3} and {L7, L4, L1} with X.

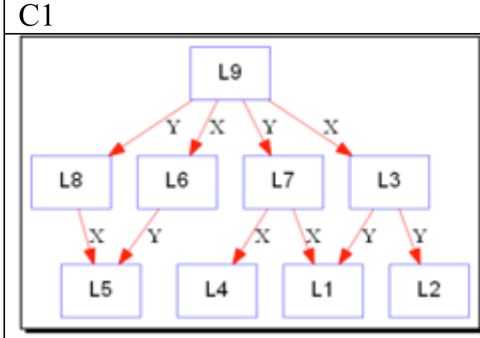
43) *$T_{2CoreXY}$*

a. *mUVT*

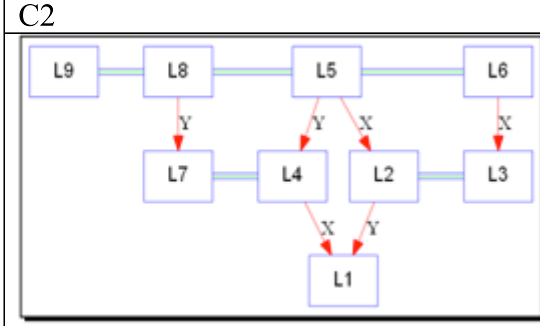
Γ	X	Y	C2	C1
L1			2	2
L2		1	2	1
L3		2	1	1
L4	1		2	1
L5	1	1	2	
L6	1	2	1	
L7	2		1	1
L8	2	1	1	
L9	2	2		

b. *EPOs*

C1



C2



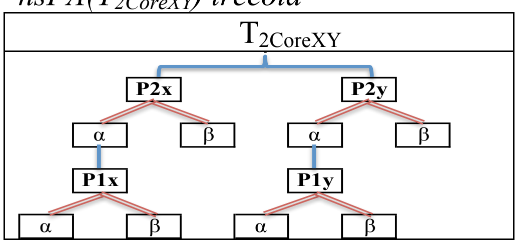
Each subPA is a PA(T_{2Core}), generating three Γs (value combinations); free combination of these values defines the 9 Γs ∈ $T_{2CoreXY}$. The wsPA and nsPA are given in parallel in (44), as the properties are familiar from T_{2Core}; Ps are subscripted by subPA antagonist.

44) *PA($T_{2CoreXY}$): Properties*

SubPA	*Ps*	*ws*	*ns*
X	$P2_X$	X <> C2	X <> C2
		α. `WeLe` / β. `LeWe`	α. `WeLe` / β. `LeWe`
	$P1_X$	X <> {C2, C1}.dom	X <> C1 /$P2_X$.α
		α. `WeLL` / β. `LeWW`	α. `WeeL` / β. `LeeW`
Y	$P2_Y$	Y <> C2	Y <> C2
		α. `eWLe` / β. `eLWe`	α. `eWLe` / β. `eLWe`
	$P1_Y$	Y <> {C2, C1}.dom	Y <> C1 /$P2_Y$.α
		α. `eWLL` / β. `eLWW`	α. `eWeL` / β. `eLeW`

The compositional relationship between $T_{2CoreXY}$ and T_{2Core} is highlighted by the nsPA treeoid. It is exactly two copies of the nsPA(T_{2Core}) treeoid, the two subPAs, joined under a root node. The scope of each P1 is defined by a single P2.

45) *nsPA($T_{2CoreXY}$) treeoid*



The values multiply out as shown in the value table. Each of the three possible combinations of values in each subPA (X-centric combinations boxed in bold) is trifurcated by those of the other. For example, $P2_X$.α+$P1_X$.α (=L1 ∈ T_{2Core}) splits into L1.1, L1.2, and L1.3 by Y subPA values.

46) *Values tables, annotated*

Γ	*ws PA*				*ns PA*				Rankings	Γ (X.Y.C2.C1)
	2_X	1_X	2_Y	1_Y	2_X	1_X	2_Y	1_Y		
L1.1	α	α	α	α	α	α	α	α	X&Y > C2&C1	`WeLL, eWLL`
L1.2	α	α	α	β	α	α	α	β	X > C1 > Y > C2	`WLLL,eLLW,eWLe`
L1.3	α	α	β	β	α	α	β		X > C2&C1; C2 > Y	`WLLL,eLWe`
L2.1	α	β	α	α	α	β	α	α	Y > C1 > X > C2	`LWLL,LeLW,WeLe`
L2.2	α	β	α	β	α	β	α	β	C1 > X&Y > C2	`LLLW,WeLe,eWLe`
L2.3	α	β	β	β	α	β	β		C1 > X > C2 > Y	`LLLW,WLLe,eLWe`
L3.1	β	β	α	α	β		α	α	Y > C2&C1; C2 > X	`LWLL,LeWe`
L3.2	β	β	α	β	β		α	β	C1 > Y > C2 > X	`LLLW,LWLe,LeWe`
L3.3	β	β	β	β	β		β		C2 > X&Y	`LLWe`

Each Γ is the union of the rankings of the stringency set relative to each antagonist, determined independently. Only in L3.3, defined by all β values, is C1 completely unranked relative to both X and Y, as C2 dominates both. In nsPA, both P1are moot in this Γ. Each P1 is moot in 3 Γs; in these, C1 is crucially ranked relative to only one of X and Y, whichever dominates C2. This same structure generalizes both to any T_{nCore}, n>2, expanding each subPA accordingly, and to systems with more antagonists, X, Y, Z…, increasing the number of subPAs.

3.3.2.2 *Stringency set antagonist: $T_{n\times m}$*

In the second variation, the antagonist is itself a set of stringency Cs, so that CON consists of two sets of *n* and *m* Cs. The structure of such systems depends on the relationship of the scales the sets are defined by. The *n*C and *m*C sets can be defined in opposing orders along the same scale. In this case, the *n*Cs become equivent, not stringent, under the filtrations by *m*Cs and vice versa. An example would be a set of markedness Cs violated by [+voi] scaled by sonority, (e.g., a C m.d = no voiced stops, m.dv = no voiced stops or fricatives, etc.), and a set of faithfulness Cs violated by changed voicing value scaled in the opposite direction (e.g., f.v = faith to fricatives, f.dv = faith to stops and fricatives).

The *n*C and *m*C sets can also realize distinct scales, as in the system below, in which case such an equivalence relation is not derived. The T in these cases, $T_{n\times m}$, cross combines Ps of a T_{nCore} and T_{mCore}. Classification of intermediate cases is beyond the scope of the present chapter.

In Alber's (2015a) full SR typology, a markedness scale is defined by sonority, and a faithfulness scale by position (initial or internal). Her full system has 3 markedness Cs, for 3 levels of sonority, and 2 faithfulness Cs, general and positional. T_{SR2} simplifies to a 2C markedness set, instantiating the most basic $T_{n\times m}$ system, where *n, m* = 2.[15] Con_{SR2} includes all of the constraints of Con_{SR}, repeated below with the added faithfulness C, f1, violated by unfaithful mappings in internal positions only. GEN includes candidates with both initial (#) and internal (_) /sc/ clusters.

47) *SR2 GEN and Con*

a. *Gen*: Inputs: /sc/, c ∈ {k, n}, k = [-sonorant], n = [+sonorant]

Outputs: {#sc, #ʃc, _sc, _ʃc}, $c_{out} = c_{in}$.

b. *Con*: m2: *{sk,sn} (m.kn)

m1: *sk (m.k)

f2: *(S_{in}, S_{out}): S_{in} = s & S_{out} = ʃ.

f1: *(S_{in}, _S_{out}): S_{in} = s & _S_{out} = ʃ.

The UVT and MOAT show that both stringency sets in this system meet the stringency definition. In the EPOs, isomorphic across the sets, each EPO(2) orders both EPO(1) equivalence classes; red boxes show the ordering in a top class, and blue in a lower.

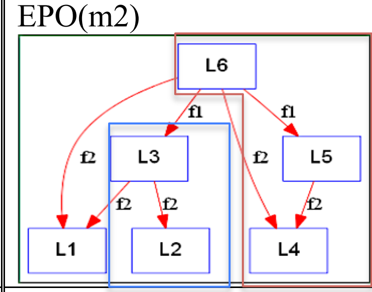
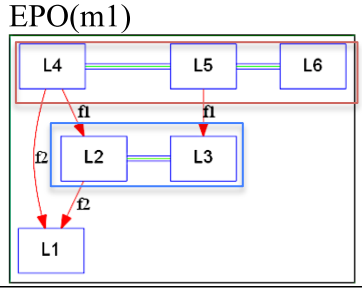
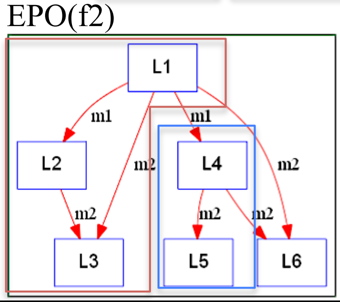
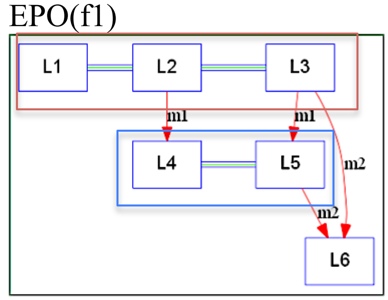
[15] For the full system, with a 3C markedness scale, see Alber's (2015a) insightful analysis. The current presentation departs from hers in the use of disjunctive scopes in nsPA, instead of repeating Ps in the treeoid.

48) *T_{SR2}*

a. *UVT*

	f2	f1	m2	m1
L1			2	2
L2	1		2	1
L3	2		1	1
L4	1	1	2	
L5	2	1	1	
L6	2	2		

b. *MOAT*

<table>
<tr><th></th><th>C2</th><th>C1</th></tr>
<tr><td>m</td><td>EPO(m2)

</td><td>EPO(m1)

</td></tr>
<tr><td>f</td><td>EPO(f2)

</td><td>EPO(f1)

</td></tr>
</table>

The wsPA cross-multiplies T_{2Core} Ps in that each stringency set antagonist of one scale, C2 and {C2, C1}.dom, is antagonized with each such antagonist for the other, resulting in four properties. As in wsPA(T_{nCore}), consistent value combinations are restricted by entailments and contradictions between the P values, generating 6 Γs (value table).

49) *wsPA(T_{SR2})*

a. *Properties*

Properties	*α*	*β*
P2.2: f2 <> m2	WeLe	LeWe
P2.1: f2 <> {m2,m1}.dom	WeLL	LeWW
P1.2: {f2,f1}.dom <> m2	WWLe	LLWe
P1.1: {f2,f1}.dom <> {m2,m1}.dom	WWLL	LLWW

b. *Value table*

Γ	P2.2	P2.1	P1.2	P1.1
L1	α	α	α	α
L2	α	β	α	α
L3	α	β	α	β
L4	β	β	α	α
L5	β	β	α	β
L6	β	β	β	β

The nsPA(T_{SR2}) antagonizes each C from the *n* set with each from the *m* set individually (50). The scopes of nsPs are defined disjunctively: either of the rankings under which a more general C is dominated for *either* stringency set. For example, P1.1 antagonizes the less stringent Cs, m1 <> f1; a Γ has a value of P1.1 if either f1 > m2 (P1.2.α) (m2 is dominated; m1 and f1 conflict) or m1 > f2 (P2.1.β) (f2 is dominated, f1 and m1 conflict). The scope of P1.1, Σ(P1.1) is P2.1.β ⋁ P1.2.α. P1.2 and P2.1 scopes are defined by single P values, as both involve one of the more general Cs as an antagonist.

50) *nsPA(T_{SR2})*

a. *Properties*

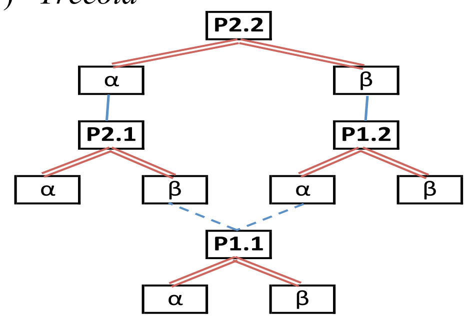
Properties	*Scope*	*α*	*β*
P2.2: f2 <> m2		WeLe	LeWe
P2.1: f2 <> m1	P2.2.α	WeeL	LeeW
P1.2: f1 <> m2	P2.2.β	eWLe	eLWe
P1.1: f1 <> m1	P2.1.β ⋁ P1.2.α	eWLe	eLWe

b. *Value table*

	P2.2	P2.1	P1.2	P1.1	*Ranking in Γ*
L1	α	α			f2 > m2 & m1
L2	α	β		α	f1 > m1 > f2 > m2
L3	α	β		β	m1 > f2 & f1; f2 > m2
L5	β		α	α	f1 > m2 & m1; m2 > f2
L5	β		α	β	m1 > f1 > m2 > f2
L6	β		β		m2 > f2 & f1

The scopes are shown in the treeoid, reflecting the overlapping nature of the Ps. There are four overlapping T_{2Core}-type treeoids embedded: P2.2 with each of P2.1 and P1.2, and each of these with P1.1. The structure is distinct from the compositional nature of nsPA($T_{2CoreXY}$), with two independent T_{2Core} structures. Dotted lines indicate disjunction in that for a P dominated by dotted lines, choice of *any* of dominating node requires choice of P value.

51) *Treeoid*



T_{SR2} is a $T_{2\times2}$; as with PA(T_{2Core}) and PA(T_{nCore}), the PA structure generalizes to PA($T_{n\times m}$) for any n and m values. The PA cross-combines Ps of PA(T_{nCore}) and PA(T_{mCore}) as above, essentially substituting each mC antagonist for X in a PA(T_{nCore}) and vice versa. This multiplies the Ps, $n\times m$ total. Formally:

$\forall(\mathrm{P}n, \mathrm{P}m)$: $\mathrm{P}n \in \mathrm{PA}(\mathrm{T}_{nCore})$, $\mathrm{P}m \in \mathrm{PA}(\mathrm{T}_{mCore})$,

$\exists \mathrm{P}n.m \in \mathrm{PA}(\mathrm{T}_{n\times m})$: $\alpha(\mathrm{P}n.m) = \beta(\mathrm{P}n)$ & $\beta(\mathrm{P}n.m) = \beta(\mathrm{P}m)$.

In a wsPA($T_{n\times m}$) each κ.dom from the n set, {A1,…,An} (red), is antagonized with each κ.dom from the m set,{B1,…, Bm} (blue) (ERC order: An.…A1|Bm.…B1).

52) *wsPA($T_{n\times m}$)*

<table>
<tr><th colspan="2">P</th><th>α</th><th>β</th></tr>
<tr><td>Pn.m</td><td>An <> Bm</td><td>Wee… | Lee…</td><td>Lee… | Wee…</td></tr>
<tr><td>Pn.$\bar{m}$</td><td>An <> {Bm, B$\bar{m}$}.dom</td><td>Wee… | LLe…</td><td>Lee… | WWe…</td></tr>
<tr><td>P$\bar{n}$.<i>m</i></td><td>{A<i>n</i>, A$\bar{n}$}.dom <> Bm</td><td>WWe… | Lee…</td><td>LLe… | Wee…</td></tr>
<tr><td>…</td><td>…</td><td>…</td><td>…</td></tr>
<tr><td>Pn.1</td><td>An <> {Bm,…B1}.dom</td><td>Wee… | LLL…</td><td>Lee… | WWW…</td></tr>
<tr><td>P1.m</td><td>{An,…A1}.dom <> Bm</td><td>WWW… | Lee…</td><td>LLL… | Wee…</td></tr>
<tr><td>…</td><td>…</td><td>…</td><td>…</td></tr>
<tr><td>P1.1</td><td>{An,…A1}.dom <> {Bm,…B1}.dom</td><td>WWW… | LLL…</td><td>LLL… | WWW…</td></tr>
</table>

In nsPA($T_{n\times m}$) (53), the scope of each nsP*x*.*y* is defined by the disjunction of Σ(P*x*) in PA(T_{nCore}) and Σ(P*y*) in PA(T_{mCore}): Σ(P*x*.*y*) = P*x*+1.*y*β ⋁ P*x*.*y*+1.α, the values that define, for each set, the rankings under which the more stringent C in that set is dominated.[16]

53) *nsPA($T_{n\times m}$)*

P		*Scope*	*α*	*β*
P*n*.*m*	A*n* <> B*m*		Wee…|Lee…	Lee…|Wee…
P*n*.$\overline{m}$	A*n* <> B$\overline{m}$	P*n*.*m*.α	Wee…|eLe…	Lee…|eWe…
P$\overline{n}$.*m*	A$\overline{n}$ <> B*m*	P*n*.*m*.β	eWe…|Lee…	eLe…|Wee…
…	…	…	…	…
P*n*.1	A*n* <> B1	P*n*.2.α	Wee…|ee…L	Lee…|ee…W
P1.*m*	A1 <> B*m*	P2.*m*.α	ee…W|eLe	ee…L|Wee…
…	…	…	…	…
P1.1	A1 <> B1	P1.2.α ⋁ P2.1.β	ee…W|ee…L	ee…L|ee…W

3.3.2.3 Multiple overlapping scales

In the systems analyzed above, all stringency-related Cs are in the same ordered set. With multiple overlapping sets, a C1a and C1b may both be less stringent than C2, but lack a stringency relationship between them, defining two distinct ordered stringency sets, C2 > C1a and C2 > C1b. In COT, this may occur when the sets realize a scale over distinct domains. For example, consider two Cs, f.V.rt, root-faithfulness to [±voi], and f.+V, faithfulness to [+voi], not [-voi]. Both may be less stringent than a general [±voi] faithfulness, f.V, but are likely in no such relation relative to each other: f.V.rt refers to morphological structure, f.+V to a feature value[17].

In the AOT system $T_{2.2core}$, there are two 2C stringency sets, sharing C2 as the more stringent in each. The mUVT and EPOs of these Cs are shown below. Each of the

[16] The generalized PA does not lend itself to an easily representable treeoid structure, because it depends on *n* and *m*, ranging from uniform branching to lattice-like. See also §3.6 on treeoidal similarities with certain WOT PAs.

[17] Whether these could end up in a derived stringency relationship depends on the particulars of the system. Prince & Tesar (2004:§6) construct an example where a stringency relation emerges between an f.σ_1 (first syllable) and f.σ' (stressed syllable) under a hierarchy.

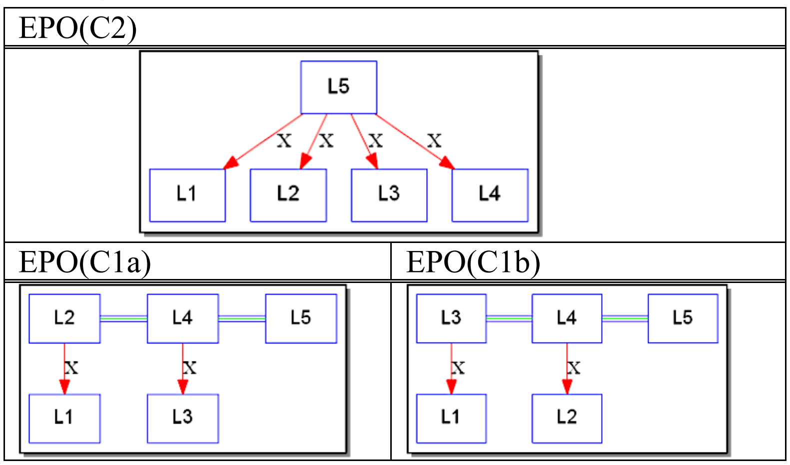
EPO(C1)s have the MOAT mark relative to EPO(C2): EPO(C2) orders the equivalence classes {L2, L4} in EPO(C1a) and {L3, L4} in EPO(C1b). The two C1 EPOs illustrate unrelated Cs.

54) $T_{2.2Core}$

a. *mUVT*

	C2	C1a	C1b	X
L1	3	1	1	
L2	2		1	1
L3	2	1		1
L4	1			2
L5				3

b. *EPOs*



PA($T_{2.2Core}$) expands PA(T_{2Core}) through replicating P1 for each C1. In wsPA, each occurs in a κ.dom with C2, but not with the other C1. In nsPA (below), the two P1s have the same scope: for either C1, its ranking relative to X depends on that of C2 and X (P2α). P1 values combine freely, generating four Γs in which X > C2. The treeoid shows the duplication: the T_{2Core} structure is refined through the lower-level split of a single P1 to separate Ps, both dominated by the same value. The structure is also a top-level collapsing of the two wsP nodes for nsPA($T_{2CoreXY}$), having a single P2.

55) *nsPA*($T_{2x2Core}$)

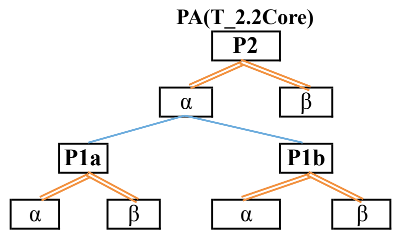
a. *Properties*

Property		*α*	*β*
P2	X <> C2	WLee	LWee
nsP1a	X <> C1a	WeLe	LeWe
nsP1b	X <> C1b	WeeL	LeeW

b. *Value table*

	P2	P1a	P1b
L1	α	α	α
L2	α	β	α
L3	α	α	β
L4	α	β	β
L5	β		

c. *Treeoid*



The variations on the core structure examined here do not exhaust the possibilities—for example, the overlapping sets in the last case could have distinct antagonists—but serve to show how the intensional structure recurs across different systems with Cs in a stringency relation. Additional cases in the appendix involve a relation holding between a C and a set of Cs. The final subsection shows how understanding these core structures facilitates analysis of COT systems. In §3.4 cases of partial stringency are analyze; while greater departures from the core structure, the main characteristics emerge, alongside those of other relations.

3.3.3 Stringency PAs in action: analyzing LingPulmAlt (LPA)

As the previous sections show, stringency systems are characterized by a fundamental set of Ps, that expands in various systematic ways. Identification of a stringency relation in a system provides a near immediate analysis and understanding of its typology. Such a strategy is illustrated here for the COT system LingPulmAlt (LPA), a modification of Bennett (2017) LingPulm system analysis. It differs from his analysis in using general faithfulness Cs, rather than separate Cs specific to features [±lingual] or [±pulm].

The system derives Bennett's insight of the cross-linguistic distribution of nasal clicks. Oral clicks are *more marked* than nasal: any language with oral clicks also has nasal, but not vice versa. Additionally, some languages contextually restrict the distribution of clicks: they are less marked word-initially than in non-initial positions (see Bennett 2017 for detailed empirical typology). GEN and CON are given in (56); following Bennett, clicks are [+ling]; click nasality is distinguished by [±pulm]. Clicks are represented orthographically by capitals, N and Q, non-clicks by lowercase k and q.

56) *LPA: GEN and CON*

a. *Segmental feature representations*

	+ling	−ling
+pulm	N	k
−pulm	Q	q

b. *GEN:* Inputs/outputs: Xa.Ya: X, Y ∈ {N, Q, k, q}

c. *CON*

m.L: *Q,N — *violated by clicks*

m.Agr.P: *Qa,qa,aQ,aq — *violated by adjacent segment* [±*pulm*] *disagreement*

f.F: $*(S_{in}, S_{out})$: $[\alpha F] \in S_{in}$ & $[\neg\alpha F] \in S_{out}$, F ∈ {[ling], [pulm]}.

f.in.F: $*(\#S_{in}, \#S_{out})$: $[\alpha F] \in S_{in}$ & $[\neg\alpha F] \in S_{out}$, F ∈ {[ling], [pulm]}

GEN produces 16 possible inputs, with 16 outputs each; 4 of these are a Universal Support[18]. These are shown below; HB candidates are removed, leaving 2 possible optima.

[18]This was established after calculating the system with the full GEN. For all other csets, there is either a single optimum, or the mapping is predictable based on those of the 4 US csets.

57) *LPA: US csets*

Input	Output	m.L	m.Agr.P	f.F	f.in.F
Naka	Naka	1			
	kaka			1	1
kaNa	kaNa	1			
	kaka			1	
qaka	qaka		1		
	kaka			1	1
kaqa	kaqa		2		
	kaka			1	

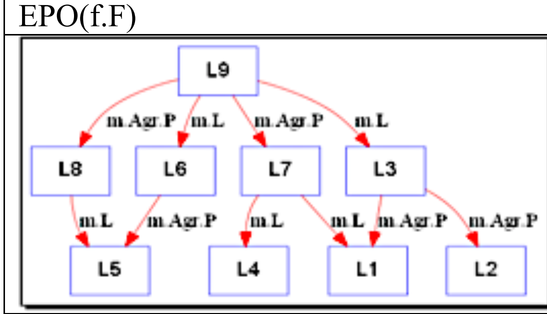
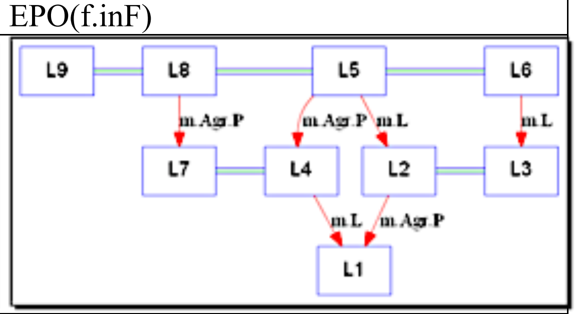
The stringency relation is identifiable from the mUVT and EPOs (58), which are equivalent to those of $T_{2CoreXY}$. Over unfiltered K, f.F[K] = L9 ⊂ f.in.F[K] = {L5, L6, L9}; subset relations also hold for h = m.L and m.Agr.P.

58) *LPA*

a. *mUVT*

Γ	m.L	m.Agr.P	f.F	f.in.F
L1			2	2
L2		1	2	1
L3		2	1	1
L4	1		2	1
L5	1	1	2	
L6	1	2	1	
L7	2		1	1
L8	2	1	1	
L9	2	2		

b. *EPOs*

EPO(f.F)	EPO(f.inF)
	

T_{LPA} exactly instantiates $T_{2CoreXY}$, with the markedness Cs, m.L and m.Agr.P, as X and Y. The full analysis follows (nsPs in (59)). The value table is extended to show the contexts in which N and q are faithful. The three choices of degree of faithfulness—none, initial # only, or all—are independently determined for each segment by the three value combinations in each subPA.

59) *PA(LPA)*

a. *Properties*

SubPA	*Ps*	
L	$P2_L$	m.L <> f.F
	$P1_L$ /$P2_{L}$.α	m.L <> f.in.F
A	$P2_A$	m.Arg.P <> f.F
	$P1_A$ /$P2_{A}$.α	m.Arg.P <> f.in.F

b. *Value table*

Γ	2_L	1_L	2_A	1_A	N	q
L1	α	α	α	α	--	--
L2	α	α	α	β	--	#q
L3	α	α	β		--	q
L4	α	β	α	α	#N	--
L5	α	β	α	β	#N	#q
L6	α	β	β		#N	q
L7	β		α	α	N	--
L8	β		α	β	N	#q
L9	β		β		N	q

The results of the PA developments in this section thus provide the basic units of analysis for any system where a stringency relationship is shown to exist. Mitchley & DelBusso (in prep.) use this strategy in analyzing a very large and complex typology with 7 Cs and 348 languages (from Mitchley 2016), by identifying a core stringency relation therein.

3.4 *Partial stringency*

In cases of partial stringency, two Cs are stringently-related under some, but not all hierarchies. Under those where they are not, other conflict or non-conflict relations between the Cs can obtain. But as stringency is somewhere present, so too are its characteristic structures, in MOATs and PAs, coexisting with the characteristic structures of other relations. This section analyzes three cases: two where stringency is 'lost' under some filtration product (T_{AP}, T_{Conf}), the last where it is 'derived' (T_{der}). All are AOT systems, constructed to isolate the relations of interest, generally resulting from simplifications of COT cases, as noted in the sections.

3.4.1 Lost: stringency + non-conflict

Prince (2000, 2001) characterized rankings in which a more stringent C dominates a less as anti-Paninian (AP). With a filtration definition, AP rankings do not occur with global stringency, but only where C2 filtration is *not* a subset of C1 filtration. For this loss of stringency to arise, there must be a hierarchy, h, such that h.C2[K] $\not\subseteq$ h.C1[K] and under h[K], both C1 and C2 conflict independently with another antagonist X. For an AP ranking, C2> … >C1, C2 dominates X and h.C2[K] must be non-decisive so that C1 and X ranking determines optimum.

In the AOT system T_{AP}, h = Y, seen in the mUVT. Under this filtration, Y[K] = {L1.1, L2.1, L3.1.1, L3.1.2} there is no subset relationship between the filtrations of C2 and C1. While L3.1.2 is in both, and L1.1 in neither, they split on L2.1 and L3.1.1, both conflicting with Y. Neither is stringent with regard to the other. As a result, T_{AP} refines $T_{2CoreXY}$ by splitting a Γ in which C1 and X are not ranked into two distinct Γs.

60) *T_{AP} mUVT*

	Y	X	C2	C1
L1.1			3	2
L2.1		1	3	1
L3.1.1		2	2	2
L3.1.2		3	2	1
L1.2	1		2	1
L1.3	2		1	1
L2.2	1	1	2	
L2.3	2	1	1	
L3.2	1	2	1	
L3.3	2	2		
Y.C[K]	--	L1.1	L3.1.1 L3.1.2	L2.1 L3.1.2

In nsPA(T_{AP}), all Ps are the same as in nsPA($T_{2CoreXY}$) (61). However, as Cs are not in a stringency relation under Y filtration, the scope of $P1_X$, ranking C1 and X, expands to the P value defining h[K]: $P2_Y.\alpha$.The scope is the disjunction of $P2_X.\alpha$ and $P2_Y.\alpha$, resulting in an additional possible value combination that splits L3.1 into L3.1.1 (a Γ_{AP}) and L3.1.2. As with unrelated Cs generally (§3.2.1), each of C1 and C2 is separately antagonized with their joint antagonist X. The treeoid (61) shows the scope relations (recall from above that a P dominated by dotted lines is non-moot under any of the dominating values).

61) *PA(T_{1AP})*

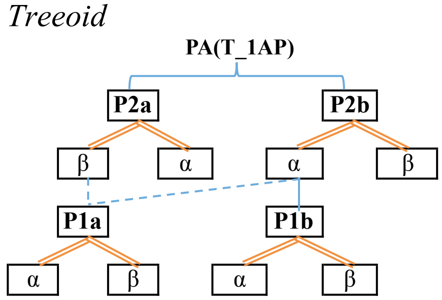
a. *Properties*

	Ps		*Scope*	α	B
X	$P2_X$	X <> C2		WeLe	LeWe
	$P1_X$	X <> C1	$P2_X.\alpha \vee P2_Y.\alpha$	WeeL	LeeW
Y	$P2_Y$	Y <> C2		eWLe	eLWe
	$P1_Y$	Y <> C1	$P2_Y.\alpha$	eWeL	eLeW

b. *Value table*

Γ	$P2_X$	$P1_X$	$P2_Y$	$P1_Y$	Rankings
L1.1	α	α	α	α	X & Y > C2 & C1
L1.2	α	α	α	β	X > C1 > Y > C2
L1.3	α	α	β		X > C2 & C1, C2 > Y
L2.1	α	β	α	α	Y > C1 > X > C2
L2.2	α	β	α	β	C1 > X & Y > C2
L2.3	α	β	β		C1 > X > C2 > Y
L3.1.1	β	α	α	α	Y > C2 > X > C1
L3.1.2	β	β	α	α	Y > C2 & C1 > X
L3.2	β	β	α	β	C1 > Y > C2 > X
L3.3	β		β		C2 > X & Y

c. *Treeoid*



A fully-wsPA is not possible: wsPA($T_{2CoreXY}$) cannot be manipulated to generate any additional Γs, as scopes cannot be widened and all unsubstantiated value combinations are inconsistent. In wsP1, domination of C1 entails domination of C2: any dominated κ.dom that includes C1 also includes C2. It is impossible to rank X and C1 to the exclusion of C2, ruling out any Γ where C2 > X > C1. To generate a Γ_{AP} nsP1(s) are necessary, moot in some Γ(s).

Further complicating, in AOT system T_{2AP}, each of Y and X acts as a stringency-losing h for the other antagonist. As with X, C1 ranking relative to Y occurs when X>C2>Y. Σ($P1_Y$) expands in a parallel way, generating an additional possible value combination, previously precluded by mootness that splits L1.3 of $T_{2CoreXY}$ by C1 and Y ranking.[19] Only when C2 dominates both antagonists—$P2_X$.β+$P2_Y$.β—are nsP1s moot.

62) *PA(T_{2AP})*

a. *Properties*

	Ps		α	β
X	$nsP1_X$	X <> C1 /$P2_X$.α ⋁ $P2_Y$.α	`WeeL`	`LeeW`
	$P2_X$	X <> C2	`WeLe`	`LeWe`
Y	$nsP1_Y$	Y <> C1 /$P2_Y$.α ⋁ $P2_X$.α	`eWeL`	`eLeW`
	$P2_Y$	Y <> C2	`eWLe`	`eLWe`

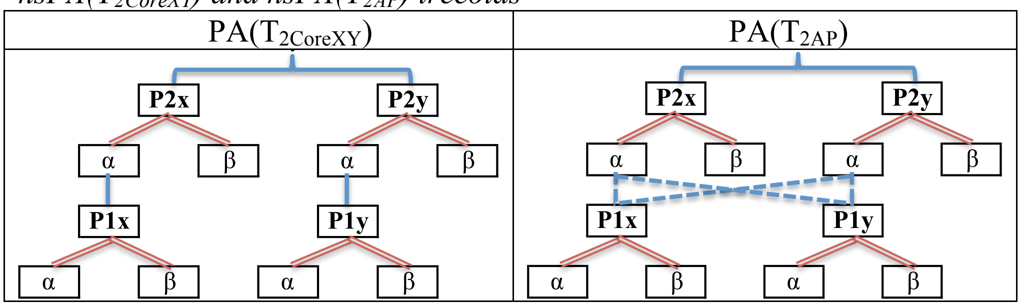
b. *Value table*

Γ	$P2_X$	$P1_X$	$P2_Y$	$P1_Y$	Rankings
L1.1	α	α	α	α	X & Y > C2 & C1
L1.2	α	α	α	β	X > C1 > Y > C2
L1.3.1	α	α	β	α	X > C2 > Y > C1
L1.3.2	α	α	β	β	X > C2 & C1 > Y
L2.1	α	β	α	α	Y > C1 > X > C2
L2.2	α	β	α	β	C1 > X & Y > C2
L2.3	α	β	β	β	C1 > X > C2 > Y
L3.1.1	β	α	α	α	Y > C2 > X > C1
L3.1.2	β	β	α	α	Y > C2 & C1 > X
L3.2	β	β	α	β	C1 > Y > C2 > X
L3.3	β		β		C2 > X & Y

[19]Other logically possible combinations are transitively contradictory, i.e., X > C2 > Y and Y > C1 > X.

The treeoids highlight the difference to the PA structure between nsPA($T_{2CoreXY}$) and nsPA(T_{2AP}). In the former, each antagonist-defined subPA involves distinct Ps; in the latter, the disjunctive scopes indicate P sharing. The result of such sharing widens the set of possible value combinations, though total free combinability is curtailed by contradictory rankings.

63) *nsPA($T_{2CoreXY}$) and nsPA(T_{2AP}) treeoids*



Partial stringency can coexist in the same system as global, for different sets of Cs (e.g. Mitchley 2016; also Mitchley & DelBusso in prep.).

3.4.2 Lost: stringency + conflict

While C2 and C1 in T_{AP} are not globally stringent, they do not conflict. Any ordering between them in Γs is by transitivity of other rankings. Two Cs can also be related both stringently and conflictingly, an apparent contradiction. However, recall that partial stringency holds for *some* filtration products only; over these, the Cs cannot conflict, by the stringent relationship, but non-conflict is not entailed for other h[K]s where no stringency exists. LVT (9) illustrates such a case: for h = ∅, the stringency-defining filtration subset relation holds; however, under a non-empty h[K], C2 and C1 conflict.

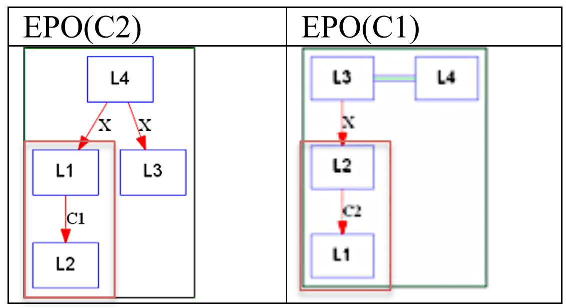
A simplification of LVT is modeled by the AOT system T_{Conf}, (UVT in (64), EPOs repeated from (27)a), which refines T_{2Core} by splitting L1, where C2 and C1 are dominated but not themselves ordered. The splittees, L1.1 and L1.2 in T_{Conf}, are the filtration product under h = X. Clearly, C1 and C2 ordering is crucial to deciding between these. C1 and C2 conflict by the definition of conflict: λ1 ∈ L1.1 = X.$\underline{C2.C1}$.Q and λ2 ∈ L1.2 = X.$\underline{C1.C2}$.Q.

64) T_{Conf}

a. *UVT*

	X	C2	C1
L1.1		1	2
L1.2		2	1
L2	1	1	
L3	2		

b. *EPOs*



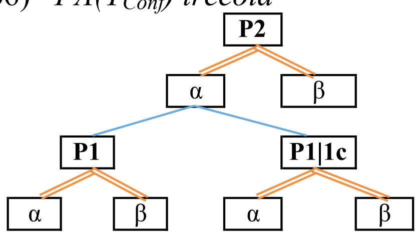
The PA shows the dual stringency + conflicting relationship between the Cs in having the hallmarks of both stringency systems (T_{2Core} Ps) and conflicting Cs (antagonists in $P1|1_C$). In T_{Conf}, h = X, so the scope of the conflict, $\Sigma(P1|1_C)$ is P2.α, X > C2; the conflict between C2 and C1 is limited to Γs in which C2 is dominated. Either wsP1 or nsP1 is possible, but $P1|1_C$ is necessarily ns.

65) *PA(T_{Conf})*

wsPA(T_{Conf})				nsPA(T_{Conf})			
P			*Scope*	*P*			*Scope*
P2: X <> C2				P2: X <> C2			
wsP1: X <> {C1, C2}.dom				nsP1: X <> C1			/P2.α
$P1\vert 1_C$: C2 <> C1			/P2.α	$P1\vert 1_C$: C2 <> C1			/P2.α
Γ	P2	wsP1	$P1\vert 1_C$	Γ	P2	nsP1	$P1\vert 1_C$
L1.1	α: WLe	α: WLL	α: eWL	L1.1	α: WLe	α: WeL	α: eWL
L1.2	α: WLe	α: WLL	β: eLW	L1.2	α: WLe	α: WeL	β: eLW
L2	α: WLe	β: LWW	β: eLW	L2	α: WLe	β: LeW	β: eLW
L3	β: LWe	β: LWW		L3	β: LWe		

The treeoid is a PA(T_{2Core}) structure, refined by the addition of P1|1c: the same central ranking choices in T_{2Core} are made in T_{Conf}, but there is an further choice in the latter. The structure is the same as that of the overlapping scales, T_{22Core} (55)c, but the content of the Ps crucially differs. In PA(T_{22Core}), each P1 antagonized X with a distinct C1; in PA(T-$_{Conf}$) the shared antagonist of P1 and P1|1c is not X but C1.

66) *PA(T_{Conf}) treeoid*



While there are four logical value combinations of P1 and P1|1_C, given their scopes, one is inconsistent: C1 and C2 are transitively ranked in L2, entailing P1|1_C.β. A P1|1_C value is not necessary to generate L2. C2 and C1 only conflict as defined in (13)—having a BPP—in L1.1 and L1.2, the filtration product under X (the sole arrow reversal in the EPOs in (27)b).

3.4.3 Derived stringency

Derived partial stringency is the reverse of lost: a stringency relation for a C1 and C2 emerges *only* under a filtration by a non-empty h, but not over unfiltered K. The characteristic MOAT and PA structures of stringency occur embedded within a larger structure. The EPOs for an AOT system modeling this, T_{der}[20], are shown in (27)a. The mUVT is below. Here, h = Y; filtration Y[K] = {L1, L2, L3} defines the scope of the C2

[20]While not abstracted from a full COT system, T_{der} was constructed after a case given in Prince & Tesar 2004:42-3.

and C1 stringency relation. Outside this scope, no such relation exists; C1 conflicts independently with {X, Y}.dom, while C2 does not.

67) *UVT: T_{der}*

	Y	X	C2	C1
L1			2	1
L2		1	2	
L3		2	1	
L4	2	1		
L5	1			1

PA(T_{def}) includes PA(T_{Core}), scoped in the h–defining value P1|1_Y.α (68). The opposite value, P1|1_Y.β, is the scope under which C1 conflicts with {X, Y}.dom. The PA(T_{2Core}) embedding is highlighted by the treeoid: PA(T_{2Core}) = {P2, P1} (boxed) occurs under P1|1_Y.α. While PA(T_{Conf}) embedded a P within PA(T_{2Core}), PA(T_{der}) embeds PA(T_{2Core}).

68) *PA(T_{der})*

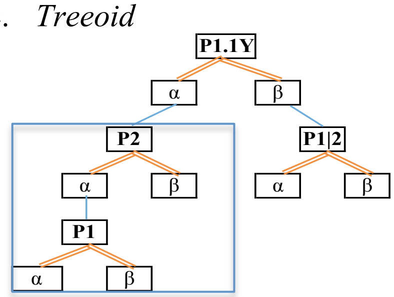
a. *Properties*

Properties	*Scope*	*α*	*β*
P1|1_Y: Y <> C2		WeLe	LeWe
P2: X <> C2	P1|1.α	eWLe	eLWe
P1: X <> C1	P2.α	eWeL	eLeW
P1|2: {X, Y}.dom <> C1	P1|1.β	WWeL	LLeW

b. *Value table*

	P1|1_Y	P2	P1	P1|2	*Ranking*
L1	α	α	α		Y > C2; X > C1, C1
L2	α	α	β		Y > C2; C1 > X > C2
L3	α	β			Y > C2 > X (C1 unranked)
L4	β			β	C2 > Y; C1 > X & Y
L5	β			α	C2 > Y; X|Y > C1

c. *Treeoid*



Though a non-exhaustive survey of kinds of C relations combinations, these cases of partial stringency show how the core structure occurs within Ts and PA, coexistent with other relations and the Ps that generate them. Other combinations are possible, using the core stringency PA pieces in different ways. For example, where Cs are equivalent and stringent, scopal inversion can occur, where the wsP has a κ.dom of stringent + equivalent Cs, while that antagonizing a C2 is ns—and that with C1 even more ns.

3.5 *Stringency PAs and extensional classification*

The languages of stringency system typologies realize steps along a scale governing the distribution of a marked extensional trait in the languages' optima. They range from having the marked trait in *all*, *none*, or *some* (aspect of) optima. How this is realized depends on the scale. For example, with a positional scale, the marked trait may be limited to a subset of environments. For a multi-point sonority-syllable peak scale, .V. $>_m$.N. $>_m$.T., a language may allow some less sonorous (N) peaks, but not the least (T).

These options correlate with P values: in a T_{nCore} system, a P*n* (X <> C*n*) makes a categorical classification of *none* vs. *some*. Each P*x*, *x*<*n*, makes the same classification over a smaller subset of cases. Uniform-value Γs (all-α/β) represent the ends of the scale, *none* (*least* marked) and *all* (*most* marked); those defined by combinations of α's and β's are the mixed *some* cases. Whether *none* correlates with α or β depends on whether the trait defined as 'marked' violates or satisfies the Cs in the stringency set.

In the next chapter, the extensional traits determined by a T_{3Core} are examined in detail. Here, the more complex case of the Alber-based S-retraction system, T_{SR2} (§3.3.2.2), shows how they play out with inter-connected scales. In this system, languages vary in the degree of SR (non-faithfulness) in their optima; lack of retraction, sc (faithfulness), is the 'marked' trait. SR distribution is scaled both by position (f Cs, internal vs. initial) and sonority (m Cs), deriving two generalizations (Alber 2015): a language may have SR in initial positions only, but no language has it in internal only; and a language may have SR before obstruents, but not sonorants only. The nsPA value table is repeated below, along with the optima for US csets; retracted [s] ([ʃ]) is shaded teal; initial and internal contexts are notated by # and _, respectively.

69) *Stringency and extensional classification: T_{SR2}*

	P values				Inputs				Classification: SR		
Γ	P2.2	P2.1	P1.2	P1.1	#sn	#sk	_sn	_sk	_k	_n	all
L1	α	α			#sn	#sk	_sn	_sk	*no*	*no*	*None*
L2	α	β		α	#sn	#ʃk	_sn	_sk	*initial*	*no*	*Some*
L3	α	β		β	#sn	#ʃk	_sn	_ʃk	*all*	*no*	
L4	β		α	α	#ʃn	#ʃk	_sn	_sk	*initial*	*initial*	
L5	β		α	β	#ʃn	#ʃk	_sn	_ʃk	*all*	*initial*	
L6	β		β		#ʃn	#ʃk	_ʃn	_ʃk	*all*	*all*	*All*

A P.β value correlates with SR; a P.α value with faithfulness (not all SR). The more αs, the more faithful (more 'marked' on the scale), more βs, the more SR (less 'marked').

70) *PA(T_{SR2}): extensional traits*

P	*SR before*	*Degree of marked trait (faithful)*	
		α	*β*
P2.2	initial c	some: [#sn]	none: all [#ʃc]
P2.1	initial obstruents	all: [#sc]	some: [#ʃk]
P1.2	internal c	some: [_sn]	none: [_ʃc]
P1.1	internal obstruents	all: [_sk]	none: [_ʃk]

The *all/some/none* set of extensional choices is the same set postulated to organize all syntactic typologies under the Parameter Hierarchies theory of Rethinking Comparative Syntax (ReCoS) project (Roberts 2010, 2012). Their proposal of typological structure is compared to PT in next chapter in the context of the word orders typology from the Final-over-Final-Condition (FOFC).

3.6 *Stringency systems and Weak Order Typologies*

Global stringency systems are a class of OT Ts sharing an intensional structure; this section compares this class with another: the Weak Order Typologies (WOTs) analyzed by DelBusso & Prince (D&P, in prep.). WOTs are characterized by different C relations, lacking stringency and having both conflict and non-stringent non-conflict. In WOTs, all Γs have isomorphic ranking structures, all permutations of CON over a particular weak orderings structure. They are named for the number of Cs in each ranking level; a 2-level WOT (2WOT), T.*n*|*m*, has *n* Cs in the top tier and *m* in the bottom. WOT structure and PAs are the subject of D&P and D (in prep.), where they are analyzed using (multiple kinds of) two distinct structures: one ws, using more κ.ops, the other ns.

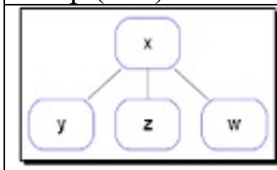
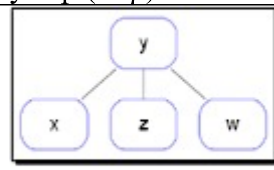
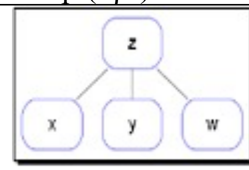
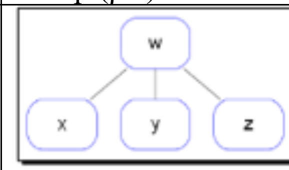
While non-equivalent classes of T, parallels arise between T.*n*|*m*s and $T_{n\times m}$s, especially in the scope structure of their (ns)PAs. These occur because of the non-rankings that occur in both: in a $T_{n\times m}$ among stringency Cs and in T.*n*|*m* among members of the same ranking level.

The 4C 2WOT T.1|3 and the core stringency $T_{1\times 3}$ are used to exemplify the PA symmetries and Γ differences, which also hold when *n*>1. Note that because $T_{1\times 3} = T_{3\times 1}$ by reversing the order of P statements, $T_{1\times 3}$ correlates equally with T.3|1, the inverse of T.1|3.[21] T.1|3 Γs are shown in (71) (P values from PA below). All Γs have a single distinct C dominating the other 3 Cs (the tops in M&P's terminology).

---

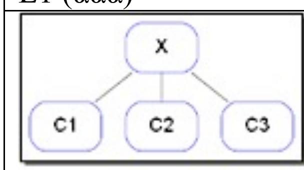
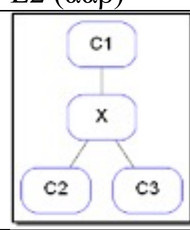
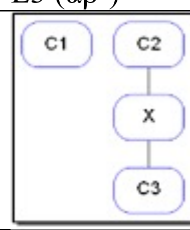
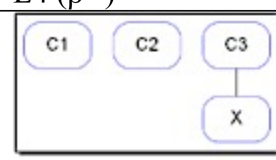
[21]PA(T.1|3) maps to PA(T.3|1) by swapping all 'dom's in PA(T.1|3) for 'sub's and α for β in value tables. This holds generally for any PA(T.*n*|*m*) and PA(T.*m*|*n*).

71) *T.1|3 Γs*

x-top (ααα)	y-top (ααβ)	z-top (αβ-)	w-top (β--)
			

The Γs in a $T_{n\times m}$ have distinct ranking structures, with some Cs are not crucially ranked (72); P values from nsPA(T_{3Core}). Only L1 is isomorphic to a $\Gamma_{T.1|3}$. L2 is a 3-level WO 1|1|2; L3 and L4 are not WOs.

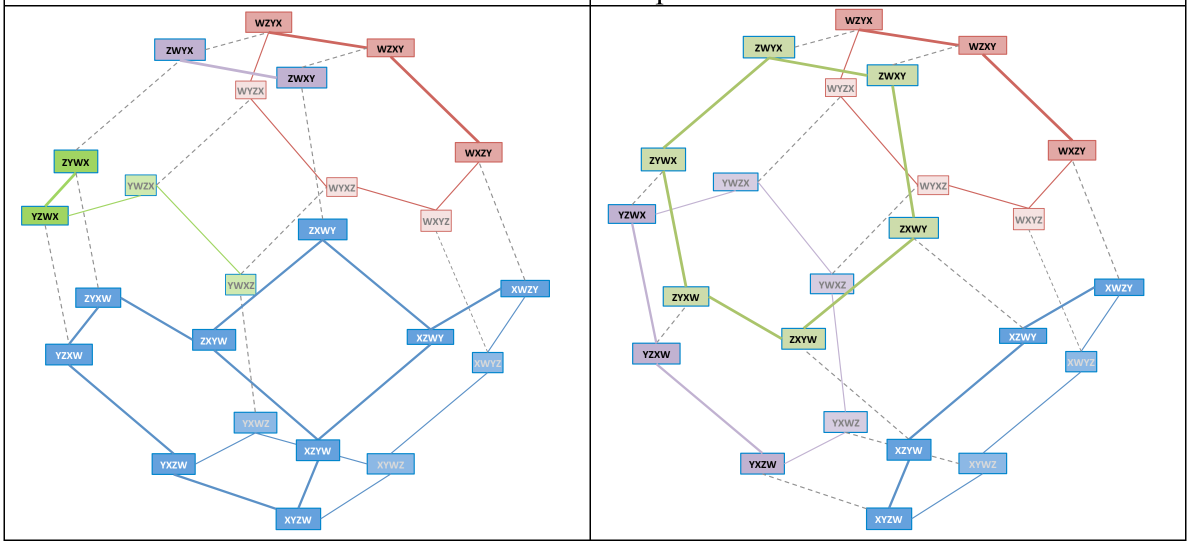
72) *$T_{1\times3}$ Γs*

L1 (ααα)	L2 (ααβ)	L3 (αβ-)	L4 (β--)
			

Mapping the Γs of each T to the 4C permutohedron (73) further highlights the differences. T.1|3 Γ all have six λs, covering a hexagonal face of the truncated octahedron. $T_{1\times3}$ Γs differ in number of λs and cover different size/shape regions: L1 is a hexagonal face (6 λ); L2 an edge (2 λ); L3 three edges (4 λ); L4 half of the λs (12) of the permutohedron.

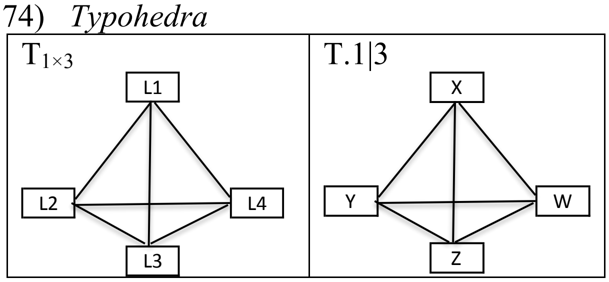
73) *Permutohedra*

T13	T.1|3
Constraints: x = C3, y = C2, z = C1, w = X *Γ*: L1 = red, L2 = purple, L3 = green, L4 = blue	*Γ*: w-top = red, y-top = purple, z-top = green, x-top = blue



Despite the non-equivalence, symmetries between the systems occur. They have the same number of Γs: $\binom{n+m}{n} = \binom{n+m}{m}$.[22] When n or m = 1, the typohedra of are isomorphic. For the example systems, n = 1, the typohedra are tetrahedra (74).

74) *Typohedra*



In both Ts, all Γs are adjacent, but for distinct reasons: in T.1|3, in each Γ the 3 dominated (non-top) Cs are unordered, with all permutations instantiated in some λ(Γ).

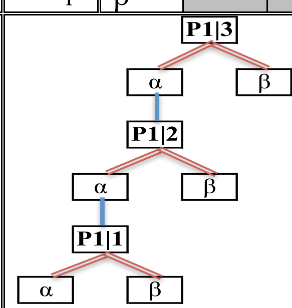
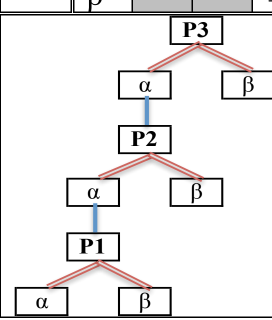
[22] See D&P on WOTs.

There is a BPP for each pair of Γs, swapping the top two Cs in λ. For example, λ_x = xyzw and λ_y = yxzw are a BPP for x-top and y-top.

Adjacency among $T_{1\times3}$ Γs arises because the locus of variation between these is ranking of the stringency Cs and X; in all but the two extremes. No two Γs differ in the ranking of more than one stringency-set C relative to X, because once a C*x* > X, all C*i*, *i*>*x* are unranked, allowing for any ordering of them in the λ. This non-ranking allows for total typohedral adjacency.

The nsPA(T_{nm}) and D&P's MA.PA(T.*n*|*m*) have the same scopal structure, highlighted by their treeoids (75), though they necessarily differ in P content.

75) *MA.PA(T.1|3) & PA($T_{1\times3}$)*

<table>
<tr><td></td><td colspan="5">MA.PA(T.1|3) (D&P)</td><td colspan="5">PA($T_{1\times3}$)</td></tr>
<tr><td>Properties</td><td colspan="5">P.1|3: P̂1|2.dom <> w
P.1|2: P̂1|1.dom <> z /P1|3α
P.1|1: y <> x /P1|2α</td><td colspan="5">P3: X <> C3
P2: X <> C2 /P3α
P1: X <> C1 /P2α</td></tr>
<tr><td rowspan="5">Value tables</td><td>Γ</td><td>P1|3</td><td>P1|2</td><td>P1|1</td><td>MIB</td><td>Γ</td><td>P3</td><td>P2</td><td>P1</td><td>MIB</td></tr>
<tr><td>x-top</td><td>α</td><td>α</td><td>α</td><td>WLLL</td><td>L1</td><td>α</td><td>α</td><td>α</td><td>WLLL</td></tr>
<tr><td>y-top</td><td>α</td><td>α</td><td>β</td><td>LWLL</td><td>L2</td><td>α</td><td>α</td><td>β</td><td>LLLW, WLLe</td></tr>
<tr><td>z-top</td><td>α</td><td>β</td><td></td><td>LLWL</td><td>L3</td><td>α</td><td>β</td><td></td><td>LLWe, WLee</td></tr>
<tr><td>w-top</td><td>β</td><td></td><td></td><td>LLLW</td><td>L4</td><td>β</td><td></td><td></td><td>LWee</td></tr>
<tr><td>Treeoids</td><td colspan="5"></td><td colspan="5"></td></tr>
</table>

In both, for each P*x* in each PA, if P*x*.β ∈ Γ, then all dependent nsPs are moot in Γ: this value indicates that all crucial rankings are accounted for. In T.1|3, this results from the use of κ.dom, ranking a 'top' C relative to all others. For example, if z > xy.dom in Γ, then x and y are not crucially ranked in λs ∈ Γ. In $T_{1\times3}$, it results from the fact that in such systems, when the more stringent C is dominant, all others in the set are not crucially ranked: if $Cx > X \in \Gamma$, then X and $\{C\overline{x},\ldots C1\}$ are not ranked in $\lambda(\Gamma)$.[23]

3.7 *Summary*

This chapter developed a definition of stringency based on filtration patterns, linking it to the formal structure of OT typologies and constraint relations. It showed how the relation is a kind of non-conflicting relation, situating it within the larger scope of kinds of C interactions. The definition and its MOAT properties also led to the identification of partial stringency, where the core interactions occur in limited domain.

PAs bring out the common intensional structure of stringency systems, occurring in all systems realizing the relationship. These results both deepen understanding of PAs and the class of stringency systems, and also provide a tool of analysis: if a stringency relation is identified within a typology (from MOAT and/or UVT scrutiny) the core properties exist in the full PA. This can be used both to yield quick grasp of simple systems, as for LPA above, or to crack more complex cases, where such Ps comprise part of the full PA.

PAs also shed further light on the extensional side of stringency. Their values precisely characterize the position on a linguistic scale, classifying the languages by the degree to which a phenomenon is manifested in it.

A. *Appendix: further aberrations of stringency*

Other variations on the common core occur when a C stands in a relation to a set of Cs. This is a common feature of Ts, and the reason for the use of κ.ops in properties: a C

---

[23] 2WOTs have several nsPAs based on different κs. MA uses uniform branching κ trees (ch. 2). Other analyses lose the symmetries with stringency $PA(T_{nm})$s.

may conflict not with each member of the class individually, but as a group. This appendix examines such cases involving a stringency relation. These are given in succinct form, with a basic description, EPOs, and PAs.

A.1. *Less stringent than a set of Cs*

Two cases are examined in which C1 is less stringent than the joint filtration of two other Cs, C2a, C2b: its filtration is either a) a superset of the intersection of the C2 filtrations: C1[K] ⊇ [C2a[K] ∩ C2b[K]][24]; or b) a superset of their union C1[K] ⊇ [C2.1[K] ∪ C2.2[K]]. The mUVTs, with filtrations for when h = ∅ in the final row, are below[25].

76) *Intersection & Union stringency mUVTs*

Int	X	C2a	C2b	C1	Un	X	C2a	C2b	C1
L1		1	1	1	L1		1	1	1
L2	1	1	1		L2	1	1	1	
L3	1	1		1	L3	2		1	
L4	2	1			L4	2	1		
L5	1		1	1	L5	3			
L6	2		1						
L7	3								
		{5,6,7}	{3,4,7}	{2,4,6,7}			{3,5}	{4,5}	{2,3,4,5}
		∩ = {7}					∪ = {3,4,5}		

In the intersection case, C1 is only unranked in L7. Its ranking relative to X remains contingent on the C2s, but occur when X dominates *either* of these. In the PA, the scope of nsP1 is the disjunction of the P values under which this obtains: P2a.α∨P2b.α. In the union case, C1 and X are only ranked when X dominates *both* C2s. The PA similarly changes the scope of nsP1, in this case to the conjunction of the values, P2a.α∧P2b.α.

[24]Or their sequential filtrations: h.C2b.C2a[K] or reverse.
[25]In both cases shown here, C2a and C2b are non-conflicting. This is not a necessary feature of systems having the kind of stringency described here.

The difference in the treeoidal representations is captured by the different line types for scopes: dotted individual lines for disjunctive, solid lines joined together for conjunctive.

77) *Intersection & Union stringency PAs*

a. *Properties (both)*

Property		*Value ERCs*	
		α	*β*
P2a	X <> C2a	`WLee`	`LWee`
P2b	X <> C2b	`WeLe`	`LeWe`
nsP1	X <> C1	`WeeL`	`LeeW`

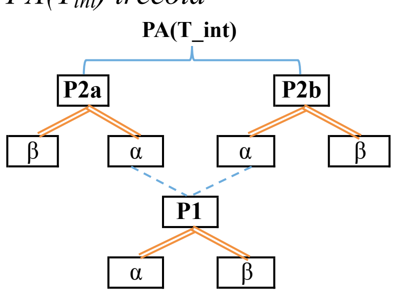
b. *PA(T_{int}) value table*

	P2a	P2b	P1
L1	α	α	α
L2	α	α	β
L3	α	β	α
L4	α	β	β
L5	β	α	α
L6	β	α	β
L7	β	β	

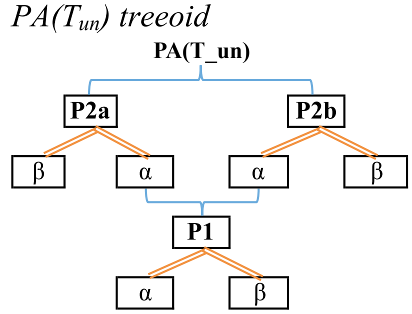
PA(T_{un}) value table

	P2a	P2b	P1
L1	α	α	α
L2	α	α	β
L3	α	β	
L4	β	α	
L5	β	β	

c. *PA(T_{int}) treeoid*



PA(T_{un}) treeoid



A.2. *Equal to a set, stringent for each*

This case holds when C2 is only more stringent than each of two C1s individually, but equal to their combination, arising when the less stringent Cs are defined on complementary subsets of C2. Prince's (2002) definition (fn 3 above) divides G's (=C2) violations between S (=C1) and D. If D is also realized as a separate C, then G's filtration is a subset of each individually, but is equal to the intersection of their filtrations. For example, if CON includes f.V, and f.+V, and f.-V. T_{S+D} (adopting Prince's notations of G = C2, S, D = C1's) is shown in UVT below.

78) *T_{S+D} mUVT*

	X	G	S	D
L1		2	1	1
L2	1	1	1	
L3	1	1		1
L4	2			

As G is equivalent to S+D (literally, its violations sum of theirs), it cannot be antagonized with X in a wsP independently of S and D. In L4, *either* G or both S and D dominate X. G occurs in a κ.dom with each of S and D in a P, with free combination of their values generating T_{S+D} (79).

79) *PA(T_{S+D})*

$P1_S$: X <> {G, S}.dom			
$P1_D$: X <> {G, D}.dom			
Γ	$P1_S$	$P1_D$	Ranking
L1	α: `WLLe`	α: `WLeL`	X > G, S, D
L2	α: `WLLe`	β: `LWeW`	D > X > G, S
L3	β: `LWWe`	α: `WLeL`	S > X > G, D
L4	β: `LWWe`	β: `LWeW`	G | S+D > X

4 The Final-Over-Final Condition and Typological Structure

4.1 *Introduction*

Linguistic theory must account for both the universals that hold of all languages and their variation within these limits. A theory generates a typology: the set of languages describable given the set of assumptions. It defines both the limits of the space of variation and the dimensions within that space on which languages can differ. Recent work under the theory of Parameter Hierarchies (Reconsidering Comparative Syntax project; ReCoS, Robert 2010, 2012, et seq.) and Property Theory (A&P, ADP) explicitly probes the internal structure of linguistic typologies, analyzing them as sets of choices with inter-dependencies limiting possible combinations. While sharing a common goal of explicating typological organization, these theories differ in significant ways.

To compare the proposals, this chapter analyzes a significant cross-linguistic generalization on possible word orders: the Final-over-Final Condition (FOFC; Biberauer, Holmberg and Roberts (BHR) 2014, Sheehan et al. 2017 and references therein). The condition expresses a gap in the typology of orders and has been a topic of much follow-up work. The current chapter focuses on the original analysis in BHR (reviewed in §4.4)[1]. This typology is also central to the development of the Parameter Hierarchy theory. The FOFC hierarchy illustrates the core aspects of the proposal: distinct parameter settings determine head-directionality of syntactic phrases in languages.

This chapter presents analyses within OT, deriving the central generalization as stated in BHR. The analysis defines a set of constraints in a stringency relationship over

[1]BHR is not the only analysis of FOFC; they discuss some alternatives in §3 of their online supplement.

positions within an Extended Projection (EP; Grimshaw 2005). The internal structure of typology is analyzed in Property Theory, explicating how the pieces of the theory generate the empirical condition. The property analyses reveal a central organizing structure into a set of properties whose values generate the precise ranking conditions aligning with the extensional trait of degree of head-finality in phrases in the language's optima. A language's property values fully determine the shape of its syntactic phrases.

FOFC follows as a consequence of the logic of OT stringency systems (chapter 3) in any system realizing the central set of stringency-related syntactic structural constraints defined over subsets of adjacent phrases within an Extended Projection. This chapter develops a set of analyses that realize this scale in distinct ways but produce intensionally equivalent systems (§4.3). OT systems with stringency constraints share a common intensional structure, regardless of the particular linguistic phenomena they explicate, as shown in chapter 3. While optima across such systems are extensionally distinct, the PAs and the logic of the explanation are the same. The property analyses further predict exactly the possible historical paths of word order change reported in BHR using Alber's (2015ab) Property Theory-based theory of diachronic variation (§4.6).

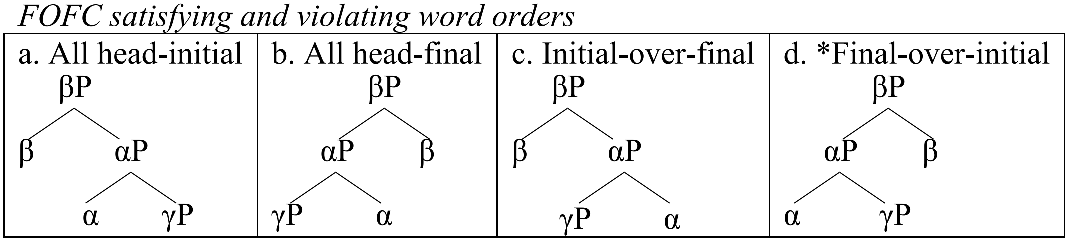
The Parameter Hierarchy proposal and Property Theory both aim to explicate typological structure more generally and use conceptually similar tools: parameters and properties (§4.5). Both structure the FOFC typology into the same range of extensional choices with crucial independencies among them. However, they differ in the structure over the choices, reflecting a deeper division in its relation to the analysis of FOFC and the source of the hierarchies. Property Theory discerns an intrinsic but non-obvious structure that is entailed by the core logic of OT. Parameter hierarchies result from a separate hypothesis, additional to and distinct from, the theoretical explanation of the analysis itself.

4.2 *FOFC and Extended Projections*

The Final-Over-Final-Condition[2] (FOFC) is a cross-linguistic generalization discovered through BHR's detailed empirical investigation. Variation in word order in syntactic phrases is restricted as in (1); structures satisfying and violating it are schematized in (2) (BHR, p. 171, (1), (2)).

1) *FOFC*: A head-final phrase βP cannot dominate a head-initial phrase αP, where α and β are heads in the same extended projection: $*[_{\beta P}\ [_{\alpha P}\ \alpha\ \gamma P]\ \beta]$.

2) *FOFC satisfying and violating word orders*



BHR characterize uniformly headed orders (2)a-b) as *harmonic* and the non-uniform (2)c-d) as *disharmonic*; however, they show that only (d) is cross-linguistically banned, based on extensive cross-linguistic study. The FOFC generalization holds for any adjacent pair of heads within the same Extended Projection (EP), and thus transitively for all heads therein. It results in a implicational statement: if β is head-final in a language, then α is, but not vice versa.

An *Extended Projection* (EP), the domain over which the condition holds, is a contiguous sequence of projections consisting of a lexical head at the base and the

---

[2]The name abbreviates the implicational statement—final in higher only if final in lower. However, a final-over-final structure is only one of the allowable structures; an alternative name is FOIC, the uniquely banned final-over-*initial* structure.

"functional shell" surrounding the lexical projection (Grimshaw 2005:2). The *categorial feature*, F, of the entire projection is inherited from the lexical category of the lexical head at the base of the EP_F, such as [+V] for verbal, [-V]/[+N] for nominal, etc. Grimshaw (2005:4 (3)) defines *head* and *projection* as follows.

3) X is a *head* of YP, YP is a *projection* of X iff:

   a. YP dominates X.

   b. The categorial features of YP and X are consistent.

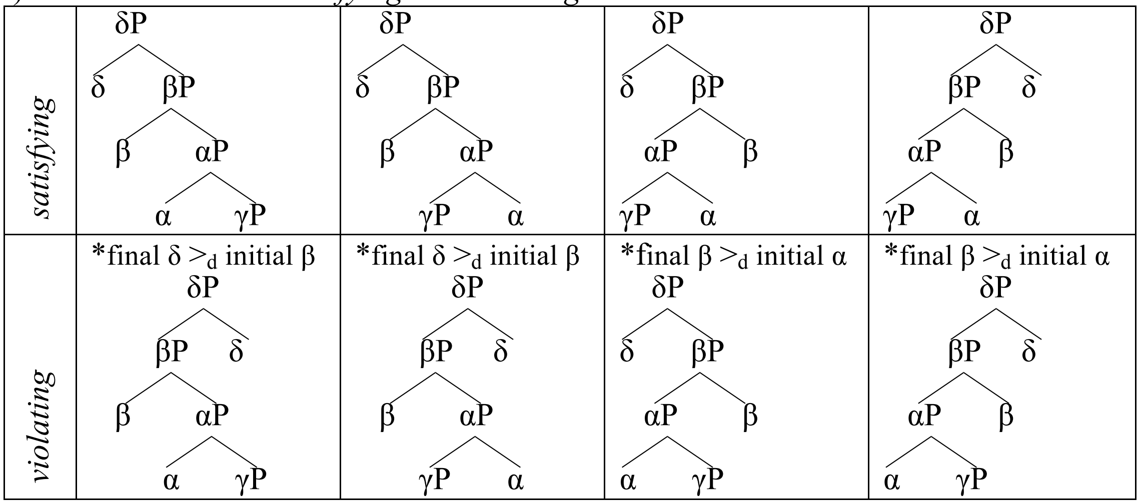
   c. There is no inconsistency in the categorial features of all nodes intervening between X and YP (where a node N *intervenes* between X and YP if YP dominates X and N and N dominates X).

Heads within an EP are ordered by their *functional value*, f-value f*n*, with the lexical head being f0, and heads above it having successively higher values. Heads are ordered in an EP such that either: a) the f-value of X is lower than the f-value of YP; or b) the f-value of X is not higher than the f-value of YP (Grimshaw 2005:4 (4)).

BHR's definition of an EP departs from that of Grimshaw, allowing a matrix clause V and a subordinate CP complement to belong to the same EP, impossible by Grimshaw's (2005) definition (BHR pp. 198-9, 211; see Biberauer & Sheehan 2012 for an analysis following Grimshaw's definition). The present chapter does not analyze the cases for which BHR require this alternative, and follows Grimshaw.

The possible orders for an EP with three heads of distinct f-value (a 3-head EP; the case examined in the analyses here) are given in (4)). The four FOFC-violating structures are marked '*' and annotated with the offending pair of heads, where $>_d$ indicates structural dominance.

4) *3-hd EP: FOFC satisfying and violating word orders*



As the number of heads in the EP increases, the number of logically possible orders of heads and complements increases exponentially (2^n for an EP with *n* f-value-distinct heads) but the number of FOFC-satisfying orders increases linearly ($n + 1$). It is this typology—the $n + 1$ FOFC word orders and none of the violating ones—that BHR and the present analyses aim to derive.

BHR derive the FOFC typology by restricting the distribution of a movement-triggering feature within an EP (§4.4). If this feature is present on a head, the complement moves to precede the head (head-finality); if not, the head precedes its complement (head-initial). BHR state "FOFC is then seen as an effect of "spreading" or inheritance of this feature from the lexical head up, from head to head within the extended projection, observing standard locality conditions on head-to-head relations" (p. 206).

FOFC is stated as a universal absolute, but potential counter-examples have been found and discussed (BHR §3, Biberauer 2017, Erlewine 2017, and references therein). The responses to such cases general fall into three main categories (adapted from Erlewine (57)): a) reject FOFC as wrong; b) show that the exception is not a counter-example, because it is not subject to the FOFC for some reason; or c) modify FOFC.

When FOFC is derived from the interaction of constraints in an OT system, exceptions receive a different explanation (see also Grimshaw 2013a on Minimalist and OT differences). In the analyses developed here using only core syntactic structure constraints, all non-FOFC candidates are harmonically bounded (HB), non-optimal under any ranking of the constraints in CON (Samek-Lodovici & Prince 2002). No exceptions are possible optima. However, harmonic bounding holds within a defined system, and can be lost when the system is modified. Non-FOFC candidates can become optimal if some other constraint(s) favoring them for a subset of cases are added, under rankings where the added constraints dominate those constraints whose satisfaction derives FOFC. Potential exception-generating constraints include: morpho-syntactic constraints that require some heads to surface as suffixes (Grimshaw, p.c.); and prosodic or discourse interface constraints that require 'light' or focused/topicalized elements to be edgemost. Exploration of this interface-exceptionality hypothesis is a topic of further research.

4.3 *Analyses: Deriving FOFC*

This section develops three OT systems analyzing the FOFC word order typology, called Sym(metric)L, Asym(metric)L, and Asym(metric)O, where the names abbreviate aspects of GEN and of CON on which they differ. All generate BHR's FOFC word-order typology by virtue of sharing the core component of a set of stringently-related constraints targeting syntactic head alignment in ordered positions within an EP. The FOFC typology follows from the logic of stringency systems. The stability of the result under these analyses underscores the crucial role of the stringency scale and also shows that it is realizable in typological equivalent systems using distinct syntactic representations and constraints.

The central component of all analyses is a set-inclusion stringency scale built on recognized structural constraints (Cs) from the literature (Grimshaw 2001 et seq.) and defined over ordered sequences of head in an EP. Cs on the scale assign violations to restricted subsets of heads or projections based on their functional values (f*n* value) in EP. The scale derives the result that the possible optimality of any candidate with head-finality in any projection depends on the order in an immediately lower or higher projection in that candidate structure.

The three variations defined below differ in which structural C the scale is built on: the alignment C Head-Left (Hd**L**, in systems Sym**L**, Asym**L**), or the obligatory-element C Obligatory-Specifier (**O**bSp, in system Asym**O**)[3]. They accordingly also differ in the antagonist to this set, a general structural C from the set {HdL, ObSp, CompL} (Grimshaw 2001).

The three typologies are both surface-order extensionally equivalent and intensionally equivalent; a system that includes a set of structural Cs scaled to heads by EP level entails the FOFC typology. That there are various possible instantiations of the necessary components replicates Bennett & DelBusso's (to appear) finding that for an Agreement-by-Correspondence (ABC) typology to produce languages with dissimilation, some correspondence C(s) in the system must have their evaluation domain restricted (by reference to features or other structures), but that restricting either type of correspondence C thusly produces the same result.

---

[3]A third option builds the scale on CompL (using $\mathrm{Gen}_{\mathrm{SymL}}$); the resulting T is equivalent to $\mathrm{T}_{\mathrm{SymL}}$, up to C relabeling.

4.3.1 The Systems: GENs and CONs

This section defines the systems, summarized below by the three dimensions of variation: i) GEN: symmetric (no movement) or antisymmetric (with movement) syntax; ii) CON: the EP-scaled C: HdL or ObSp; iii) CON: the antagonist C: HdL, ObSp, or CompL.

5) *Three systems summary*[4]

System	*SymL*	*AsymL*	*AsymO*
GEN	Symmetric	Antisymmetric	
EP scaled C	HdL		ObSp
Antagonist C	CompL	ObSp	HdL

The two GENs defined in this section differ in the base structures possible and in whether a projection in the input can appear in the different position in the output (movement). Both are simplified to generate a set of structures that vary in terms of surface order of heads and complements; other factors are held constant. The alternatives generate distinct structures for head-final orders. In GEN$_{Sym}$, either order of head and complement is possible in a projection (symmetric syntax); in GEN$_{Asym}$, phrases a strictly right-branching, with head-final order resulting from movement of the complement (antisymmetric syntax).

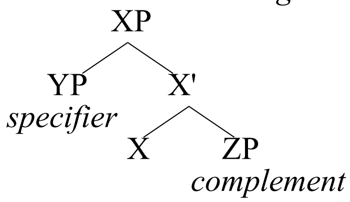
The definitions use standard terminology of syntactic X-bar structure: a *specifier* (spec) of projection XP is a maximal projection, YP, sister to X'; a *complement* (comp) is a maximal projection, ZP, sister to head X (6).

---

[4] The other 9 of 12 logically possible combinations on these criteria are eliminated as follows:

- 4: same C type (HdL or ObSp) as both the scale and the antagonist (no conflict).
- 3: Gen$_{SymL}$ + ObSp (scale or antagonist): ObSp cannot be satisfied by movement of a complement; it is equally violated by all candidates (satisfaction requires insertion of some kind, not allowed by Gen$_{Sym}$ nor relevant to the word order variation under analysis).
- 2: Asym + CompL antagonist: with Gen$_{Asym}$, CompL is filtrationally-equivalent to the most stringent HdL; it cannot be an antagonist to a HdL scale; antagonized to an ObSp scale, the T is equivalent to T$_{AsymO}$.

6) *X-bar structural categories*



Throughout, fx represents the head of projection fxP, where fx is a head in the EP with f-value x. This generalizes across EPs with distinct lexical features. It is further assumed that the identity of heads are fixed; for example, in an EP_{+V}, f0 is V, and other heads such as T have a fixed x value. ZP and YP are specifiers or complements with lowercase (xp/yp) indicating a structurally lower copy of a moved projection.

The GENs are defined in (7). Inputs for both consist of an EP_F, a fixed set of ordered heads, (f0,…,fn) and a complement ZP in a distinct EP_G, where n may vary by F. Whether all such heads have an overt lexical realization in a given language is not the subject of analysis here and would be controlled by a different constraint set. For both GENs, the complement, ZP, is treated as an unanalyzed unit, with no ZP-internal violations assessed. As the complement of a lexical head, f0, in a distinct EP, it does not incur violations of the EP-specific stringency Cs introduced below. GEN variations differ in the possible output structures, as defined below.

7) GEN:

Input: an EP_F, a set of ordered heads, (f0,…,fn) & complement ZP ∈ EP_G.

Outputs: an binary syntactic structure containing all input elements where:

a. GEN_{Sym}: ∀fx ∈ EP_F, ∃fxP ∈ out: fxP = $[_{\text{fxP}}$ fx YP$]$ or $[_{\text{fxP}}$ YP f$x]$, YP =

$$\begin{cases} \text{ZP}, x = 0 \\ \text{f}_{x-1}\text{P}, x > 0 \end{cases}.$$

- Prose: For each head, fx, in input EP_F, there is a projection, fxP, in the output, where fxP dominates head fx and its complement, YP, in either order; heads are ordered in EP_F by f-value.

b. GEN$_{Asym}$: $\forall fx \in EP_F, \exists fxP \in out: fxP = [_{fxP} \_ [_{fx'} fx\ YP]]$ or $[_{fxP}\ YP\ [_{fx'} fx\ yp]]$, YP $= \begin{cases} \mathrm{ZP}, x = 0 \\ \mathrm{f}_{x-1}\mathrm{P}, x > 0 \end{cases}$.

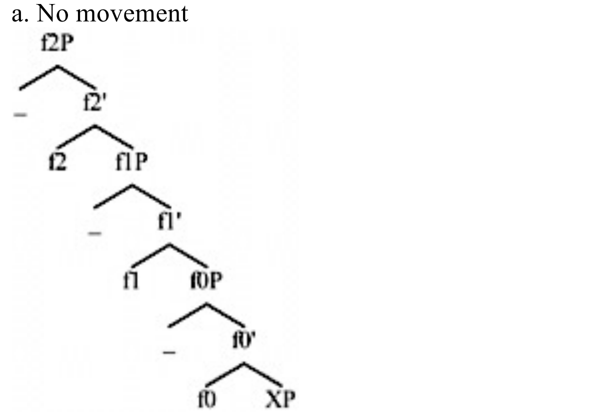
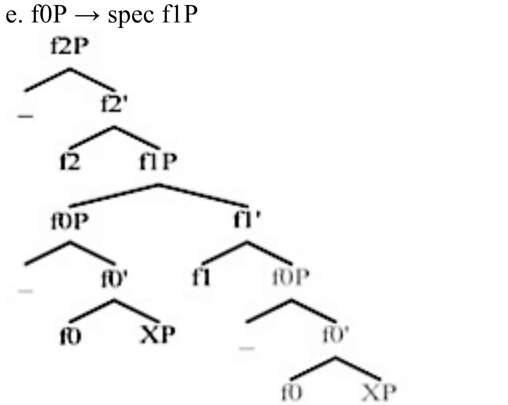
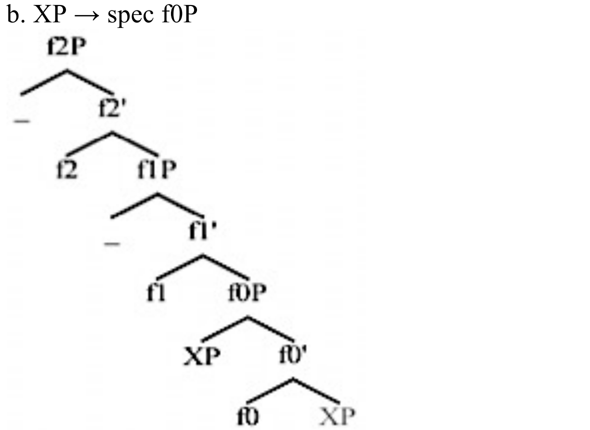
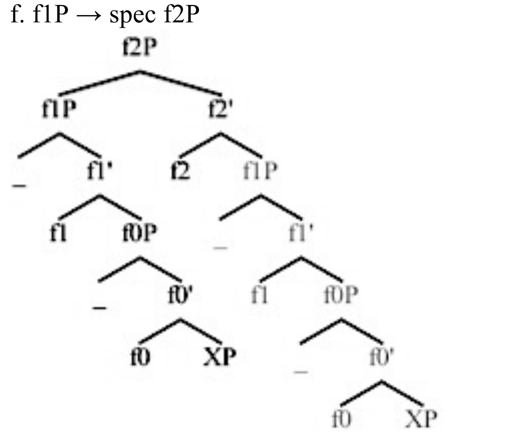
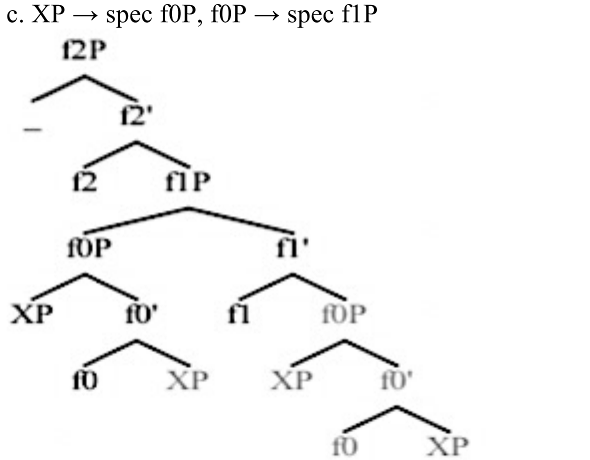
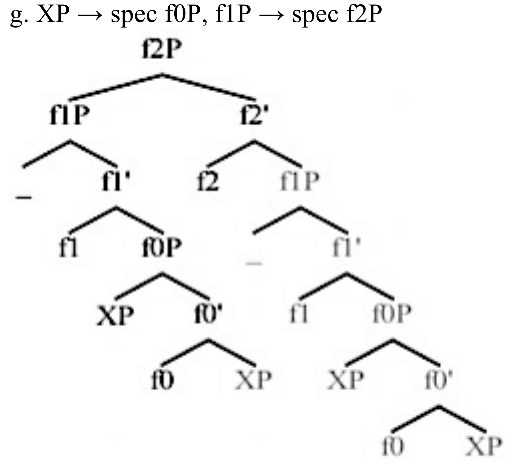
- Prose: For each head, fx, in input EP_F, there is a projection, fxP, in the output, where fxP dominates fx' that dominates head fx and comp YP, in that order, with or without a copy of comp in spec fxP; heads are ordered in EP_F by f-value.

Both GENs produce two distinct structures for any fxP, differing in the relative order of a head and complement. For GEN$_{Sym}$, either base order is possible in a projection. All outputs lack specifiers, as variation on this dimension is not relevant to head-initiality/finality in this system. A full candidate set (cset) for an EP with *n* distinct f-value heads, an *n*-hd EP, includes an output realizing each combination of the two orderings for all projections in the EP; 2^n candidates in each cset. For a 3-hd EP, the outputs are the eight structures in (4) above.

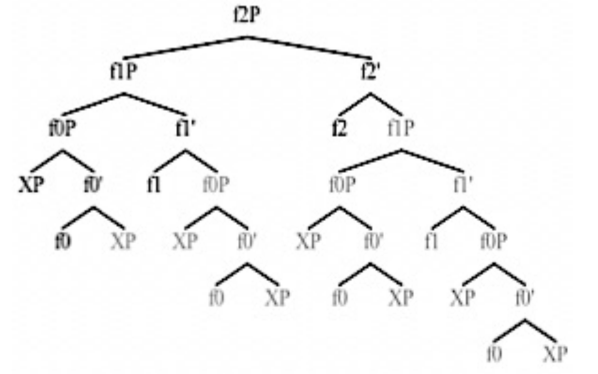
For GEN$_{Asym}$, all projections have right-branching $[_{XP}$ (spec) $[_{X'}$ X comp]] structures, where heads precede complements in a projection; head-final word order results from comp-to-spec movement, with a copy of a complement projection, f_{x-1}P moving in the specifier of its sister head, Spec fxP. In this, GEN$_{Asym}$ follows BHR's adoption of Kayne's (1994) antisymmetric syntax proposal of uniform underlying right-branching structure. The choice of presence or absence of movement for each fxP results again in 2^n candidates in each cset. The eight possible output structures for a 3-hd EP are shown in

(8): lower copies of moved f*x*P's are in grey font; '_' indicates an unfilled specifier. Outputs (a)-(d) satisfy FOFC, (e)-(h) violate it. In (e), for example, the lowest head, f0P moves to spec f1P, but f0's complement, XP, does not move to spec f0P, resulting in the banned f1-final-over-f0-initial order.

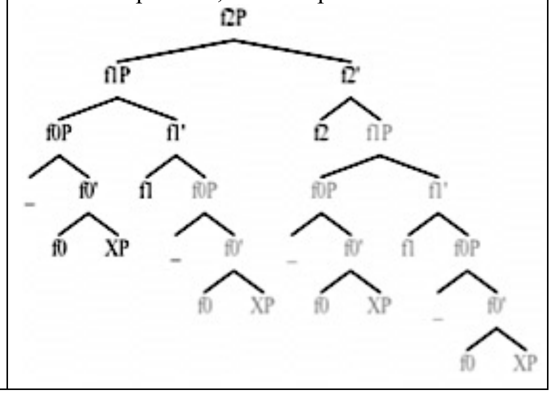
8) *GEN$_{Asym}$ 3-hd EP outputs*

a. No movement 	e. f0P → spec f1P 
b. XP → spec f0P 	f. f1P → spec f2P 
c. XP → spec f0P, f0P → spec f1P 	g. XP → spec f0P, f1P → spec f2P 

d. XP → spec f0P, f0P → spec f1P, f1P → spec f2P



h. f0P → spec f1P, f1P → spec f2P



GEN$_{Asym}$ adopts a copy theory of movement (Chomsky 1995); a lower copy of a moved projection remains in the original complement position. More movement thus results in more (copies of) projections, which has ramifications for C violation counts. This work follows Grimshaw (2001:23) in that "a moved XP and the trace of a moved XP are exactly the same, and exactly like other XPs with respect to the constraints."[5] Structural constraints are insensitive to the (un)pronounced distinction: all copies incur the same projection-internal violations, satisfy obligatory element Cs, and count as interveners for alignment of other categories. As Grimshaw shows, "since any XP incurs violations of the set of alignment and obligatory element constraints, the more occurrences of a given XP there are in a structure, the more violations there will be" (ibid.).

GEN$_{Asym}$ restricts movement to successive 'roll-up' movement (see i.e. Cinque 2005 on kinds of movement): a complement moves to the specifier position of the *same* projection, not to that of any higher projection. Spec-to-spec movement, where the specifier of a complement moves alone to a higher spec position, is excluded. In this, GEN$_{Asym}$ follows BHR, who define their movement-triggering feature as specifically

[5]Grimshaw (2001) shows that this derives economy of structure and movement.

resulting in movement of the sister of a head (the complement) to its specifier, when associated with a categorial feature F (p. 210).[6]

CON includes three structural constraints from Grimshaw (2001): two projection-internal left-alignment Cs (HdL, CompL) and one obligatory element C (ObSp)[7]. These are violated, respectively, by misalignment between the specific element with the left edge of its projection, and by the absence of a specifier within a projection. The general (non-scaled) versions are defined in (9), using the notation: f = head, ZP = comp of f; XP = any maximal projection (comp/spec). The alignment C definitions follow Hyde (2012).

9) *CON: General structural constraints*

C	*Definition*	*Prose description: one violation for each:*
HdL	*(XP, f) ∈ fP: $[_{fP}$ XP…f	(XP, f) pair in fP, such that XP intervenes between f and the left edge of fP.
CompL	*(ZP, {f, XP}) ∈ fP: $[_{fP}$ {f, XP}…ZP	(ZP, f/XP) pair in fP, such that f/XP intervenes between ZP and the left edge of fP.
ObSp	*fP: ∄XP ∈ fP: $[_{fP}$ XP f']	fP such that there is no XP, a sister of f', in fP (i.e., fP lacks a spec).

Stringency scales are defined over two of these structural constraints, HdL and ObSp. They are constructed using a subset-inclusion schema, where the set of structures to which a less stringent C assigns violations is a subset of those to which the more stringent assigns violations (Prince 2000, chapter 3 of this text). The scale references sets of heads (for HdL) or projections (for ObSp) in an EP_F, by their f-values.

[6]An alternative system, where $GEN_{AltAsym}$ includes spec-to-spec movement candidates, produces an extensionally distinct typology: the FOFC-satisfying all-final candidate is harmonically bonded by one with successive spec-to-spec movement of the lowest f0P. In this candidate, $[_{CP}$ $[_{AP}$ X A x] C $[_{BP}$ $[_{AP}$ x a x] B $[_{AP}$ x a x]]], surface order XACB, no adjacent pair of heads violates FOFC, but it is not among the structures BHR discuss as satisfying FOFC. The order is only derivable by spec-to-spec movement. A brute-force way to generate the desired FOFC T with $Gen_{AltAsym}$ uses ObSp Cs sensitive to *what* projection fills spec; specifically, satisfied only by an f*x*-1P (comp) in spec f*x*P (comp>spec).

[7]Grimshaw (2001) defines two additional structural Cs: i) ObHd (incorporated into Gen in these analyses); ii) SpecL (satisfied in all candidates in both Gens, by lack of specifiers (SymL), or by fixed antisymmetric structure, where all specifiers are leftmost in their XP (AsymL/O)).

In the HdL scale, the set of heads picked out is a contiguous sequence in an EP that includes the highest head, fn, to the lower bound indicated in the name. The most stringent, HdL.Ff0, is violated by misalignment of *any* head in the EP, f0 to fn; the least stringent, HdL.Ffn, is violated only by the misalignment of the highest head in the EP The ObSp scale works in the opposite direction: all include the lowest head, f0, up to the higher bound indicate in the name. The least stringent, ObSp.Ff0 assigns a violation to a spec-less f0P only, the *lowest* in EP_F, and the most stringent, ObSp.Ffn, to any such fxP.

The scales thus progressively isolate either the lowest or highest head in an EP, both salient elements of syntactic phrases: the lowest head is generally lexical, contributing the categorial feature of the entire EP; the highest head defines the edge of the EP. The location-specific constraints of Grimshaw (2006) target this edge position in a CP. No mid-level head is uniquely picked out. Definitions of the C scales are given in (10); Cs names pick out both F and the fx that is the *lower* (HdL) or *upper* (ObSp) bound on the set of heads.

10) *CON: Stringency scales*

C	*Definition:*	*Prose: one violation for each:*
HdL.Ffx	$\forall$f$i \in EP_F$: $x \le i \le n$, *(XP, fi) $\in$ fiP: $[_{fiP}$ XP…fi	head fi violating HdL such that fi's f-value is *greater* than or equal to x.
ObSp.Ffx	$\forall$f$i \in EP_F$: $x \ge i \ge 0$, *fiP: $\nexists$XP $\in$ fiP: $[_{fiP}$ XP fi']	projection fiP violating ObSp such that fi's f-value is *less* than or equal to x.

The definitions generate sets of Cs, one for each head in an EP_F. The set is bounded by the number of functional heads for a given F (not just those visible in a given language or input), here fixed in GEN. There are n violation-distinct Cs in the scale needed to determine optima in an n-hd EP_F input. The systems shown here use a 3-hd input EP_F, generating the 3C scale in (11).

11) *3C scales, n = 2*

Stringent	*HdL.Ffx:* *[$_{fxP}$ XP…fx	*ObSp.Ffx:* *fxP: ∄XP ∈ fxP: [$_{fP}$ {XP, fx}]
most	HdL.Ff0: x = {0,1,2}	ObSp.Ff2: x = {2,1,0}
↓	HdL.Ff1: x = {1,2}	ObSp.Ff1: x = {1,0}
least	HdL.Ff2: x = {2}	ObSp.Ff0: x = {0}

Keying the scales to the categorial feature F makes two predictions for the possible combinations of order structures within a language, aligning with the empirical generalizations BHR report.

First, it is entailed that all EPs with the same F have the same word order in all optima of the language. For example, any two EP_{+V}s have the same order regardless of where they occur in the entire structure (matrix or subordinate clause). Cs assess all such EPs equally.

Second, it is *not* entailed that any two EPs with distinct Fs have the same order in the language's optima. For example, order in an EP_{+V} may differ from that in an EP_{-V}, since each F-distinct EP is assessed by a distinct set of stringency Cs. These may be ranked relative to an antagonist independently of the ranking of any other set. The sets define *subPAs* (ch 3, also Bennett & DelBusso to appear) of the typology, and the full PA is the cross-product of the possible value combinations in each subsystem. Numerics: there are n + 1 possible optima/value combinations in a subPA with n-hd EP_F input, CON = nC scale + antagonist; for two such subPAs, input n-hd EP_F (nCs), m-hd EP_G (mCs), $T_{n×m}$ = (n + 1) × (m + 1).

4.3.2 Typologies and Property Analyses

The typologies of the three systems are calculated using a 3-hd EP_F, and the corresponding 3C stringency scale[8]. For concreteness, the input is represented as CP, with set of heads {V (f0), T (f1), C (f2)}, and complement ZP = O; the analysis extends to any 3-hd EP via relabeling. An *n*-hd input EP requires an *n*C scale; the resulting typology, T_S, has *n*+1 grammars, the structure of which predictable based on the results in the previous chapter on stringency systems.

4.3.2.1 SymL

SymL is summarized in the table repeated from above. All word orders are base-generated (no movement), with either order of head and complement possible. Head-direction in optima is determined by conflicting HdL and CompL Cs, with the stringency scale defined over HdL.

12) *SymL*

System	***SymL***	*AsymL*	*AsymO*
GEN	**Sym**	Asym	
EP scaled C	**HdL**		ObSp
Antagonist C	**CompL**	ObSp	HdL

The VT for SymL is shown in (13), with candidates represented linearly in bracket notation. Half satisfy FOFC; all that violate it are harmonically bounded, shaded in gray with the bounder(s) recorded in the final column. Any candidate with comp-head order in

[8]A Universal Support (US; Alber, DelBusso & Prince 2016) for any EP_F, n=2 input. A US for all possible phrases requires an input *n*-hd EP_F for each possible F, n = the maximum f-value in EP_F (requiring a theory of possible F's and f-values). The structure of the typology is predictable: for *m* = the number of distinct F's, n_F = the number of functional levels in the EP_F, T is the product of the *m* subsystems, where each has n_F +1 possible optima.

any projection incurs a HdL.Vf0 violation; only those with comp-head order in the highest projection, CP, incur a HdL.Vf2 violation.[9]

13) *SymL VT*

Input	Output	HdL.Vf0	HdL.Vf1	HdL.Vf2	CompL	HB-er
CP	a. [C [T [V O]]]				3	
	b. [C [T [O V]]]	1			2	
	c. [C [[V O] T]]	1	1		2	b
	d. [[T [V O]] C]	1	1	1	2	b (& c)
	e. [C [[O V] T]]	2	1		1	
	f. [[T [O V]] C]	2	1	1	1	e
	g. [[[V O] T] C]	2	2	1	1	e (& f)
	h. [[[O V] T] C]	3	2	1		

The extensional languages differ in the number of head-final projections in their optima, ranging from all-initial (L1) to all-final (L4), with the two FOFC-permitted disharmonic orders realizing steps between these extremes (L2, L3). The extensional trait of head-finality correlates with the intensional ranking of CompL dominating a subset of the HdL Cs. All languages and grammars (Γs), with their legs (λs) counts are shown in (14).[10] Constraint order in ERCs follows that in the VT. The example languages are taken from Biberauer & Roberts (2013:33).

14) *Languages and Grammars of T_{SymL}*

	Languages: optima	*Example language*	*Grammar (MIB)*	*# λ*
L1	Hd-initial: [C [T [V O]]]	English	`WeeL`	12
L2	V-final: [C [T [O V]]]	Mande (some)	`LWeL, LeeW`	4
L3	V-/T-final: [C [[O V] T]]	German	`LLWL, LLeW`	2
L4	Hd-final: [[[O V] T] C]	Japanese	`LLLW`	6

The system entails the FOFC. While head-finality can occur in any number of projections, 0 to 3, it cannot do so freely: if only one projection in optima features such an order, then it must be the lowest; if two, then the two lowest, etc. The following

[9]Following from chapter 3, this is identifiable as a T_{3Core}.
[10]See Appendix A.3 for typohedron, permutohedron, and further details of Γs.

section develops the PAs showing how the interactions of Cs ∈ CON derive the FOFC result.

4.3.2.1.1 Property Analyses

T_{SymL} has the characteristic structure of stringency systems: each C in the set of stringency scale Cs is crucially ranked relative to antagonist CompL *only* if the immediately more stringent C is dominated. As chapter 3 shows, there are alternative PAs, differing in whether all Ps are wide-scope (wsPA), using C-classes, or whether some are narrow-scope (nsPA), moot in some Γs.

The wide-scope (ws) PA properties (Ps) are stated in (15), named for the f-value of the least stringent HdL.Vf*x* in its antagonist set. In each, CompL is antagonized with a class, κ, of HdL Cs, with the operator *dom* that picks out the *dominant* member of the set in a linear order (A&P; current chapter 2). Only the most stringent, HdL.Vf0, is individually antagonized with CompL. Each less stringent C is in a κ.dom with all more stringent.

15) *wsPA(T_{SymL}): Properties*

Property		α	β
P0	CompL<>HdL.Vf0	LeeW	WeeL
P1	CompL<>{HdL.Vf0, HdL.Vf1}.dom	LLeW	WWeL
P2	CompL<>{HdL.Vf0, HdL.Vf1, HdL.Vf2}.dom	LLLW	WWWL

P value ERCs are cross-entailing. Any ERC for a P value with a dominant κ_n.dom (P*n*β) entails P(*n*+1)β, where $\kappa_n \subset \kappa_{n+1}$: P0β → P1β → P2β (by W-extension; Prince 2002). Entailment goes in the opposite direction for the reverse values (α), where κ_n.dom is subordinated, resulting in L's for each C ∈ κ: P2α → P1α → P0α (by L-retraction). As a consequence of the entailments, free combination of wsP values does not result in Γs for all 8 logical combinations (3 Ps, 2 values = 2^3); only 4 define Γs, shown in the value table

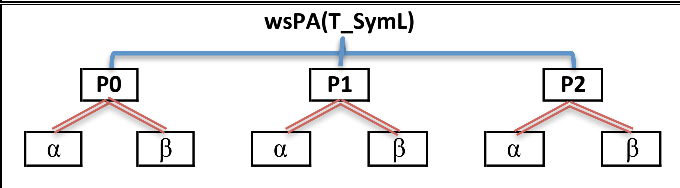
(16), with the treeoid. The value ERC sets of all others are inconsistent, a subset fusing to L+ (Brasoveanu & Prince 2011).

16) *wsPA(T_{SymL}) value table & treeoid*

a. Value table

	P0	P1	P2
L1	β	β	β
L2	α	β	β
L3	α	α	β
L4	α	α	α

b. Treeoid



To show how values derive Γs, the value ERCs and resulting MIB for L3 are shown in (17). The optima in this language realize the phrasal orders found in German.

17) *L3 in wsPA(T_{SymL})*

	value ERC	*Ranking*	*Trait*
P0α	LeeW	CompL > HdL.Ff0	hd-final VP
P1α	LLeW	CompL > HdL.Ff0 & HdL.Ff1	hd-final TP
P2β	WWWL	HdL.Ff0 | HdL.Ff1 | HdL.Ff2 > CompL	hd-initial CP
MIB	LLWL LLeW	HdL.Ff2 > CompL > HdL.Ff0 & HdL.Ff1	[C [[O V] T]]

The narrow-scope (ns) PA differs from its ws counterpart in that each C in the stringency set is individually antagonized with CompL in a P. Since only the most stringent C is crucially ranked relative to CompL in all Γs, Ps with less stringent C antagonists are ns. Their scope is defined by the P value in which the next more stringent C is dominated: Σ(P*x*) = P(*x*-1)α. Ps, value table, and treeoid showing scope structure are shown in (18).

18) *nsPA(T_{SymL})*

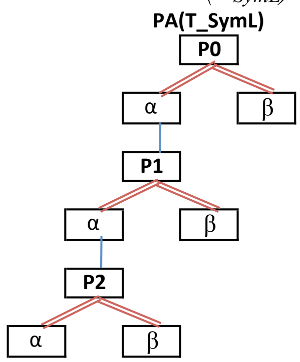
a. *Properties*

Property		Scope	α	β
P0	CompL<>HdL.Vf0		LeeW	WeeL
P1	CompL<>HdL.Vf1	/P0α	eLeW	eWeL
P2	CompL<>HdL.Vf2	/P1α	eeLW	eeWL

b. *Value table*

	P0	P1	P2
L1	β		
L2	α	β	
L3	α	α	β
L4	α	α	α

c. *Treeoid: nsPA(T_{SymL})*



The nsPA results in the same number of Γs (the viable value combinations) by mootness rather than contradiction as in wsPA: there are only four possible combinations given scopes, and all produce consistent ERC sets.

4.3.2.1.2 Deriving FOFC

FOFC is entailed in SymL: optima have head-final order in a given projection only if it occurs in all lower projections, following from the EP-based scaled HdL Cs. Head-finality is driven by satisfaction of CompL; head-initiality by satisfaction of the HdL Cs. Which of the HdLs are violated depends on which projections in a candidate have head-final order.

Degree of head-finality in the entire EP_F correlates the with P values. Head-finality occurs in a continuous sequence of the lowest *x* projections when the Γ has *x* P.α values, under which *x* HdL Cs are dominated. The table below gives the extensional traits that correlate with the P values of PA(T_{SymL}). If head-finality is, as BHR suggest, the 'marked' option, then the greater the number of α values defining a Γ, the more marked the optima in the language—more final heads. Trait and value alignment is shown in value tables of both PAs in (20), repeated from above.

19) *P values and extensional traits*

Property	*Extensional trait*	
	α: head-*final* in:	β: head-*initial* in:
P0:	f0P: …OV	f0P (f1P, f2P): CTVO
P1	(f0P) f1P: …OVT	f1P (f2P): CT[VP]
P2	(f0P, f1P) f2P: OVTC	f2P: C[TP]

20) *SymL value table and extensional traits*

a. *wsPA*

	P0	P1	P2	V (f0)	T (f1)	C (f2)
L1	β	β	β	initial	initial	initial
L2	α	β	β	final	initial	initial
L3	α	α	β	final	final	initial
L4	α	α	α	final	final	final

b. *nsPA*

	P0	P1	P2	V (f0)	T (f1)	C (f2)
L1	β			initial (all)		
L2	α	β		final	initial	
L3	α	α	β	final	final	initial
L4	α	α	α	final	final	final (all)

Two implicational generalizations hold:

1) if head f*x* is final all lower heads are final;

2) if f*x* is initial than all higher heads are initial.

These follow from the logic of stringency systems (Prince 2000, ch 3). For the wsPA, a P*n*.α (head-final) entails P*i*.α for all P*i*, $i < n$, so that if head f*n* is final, all lower heads are too. For the nsPA, the implications follow from the scope of the Ps, establishing a contingent relationship between them. A Γ only has a P(*n*+1) value if it has P*n*.α. Under P*n*.β, HdL.Ff*n* dominates CompL, and its satisfaction entails satisfaction of all HdL.Ff*x*, $x>n$, that pick out subsets of the heads picked out by HdL.Ff*n*. In wsPA, implications follow from ERC logic; in nsPA, from both scope and the filtration subset relationship of stringency Cs (chapter 3).

21) *FOFC derivation*

Rankings for finality:

- Head fx is *initial* in optima in Γ if ∀λ(Γ), HdL.Ffx ≫ CompL.
    - ∀q∈K, if fx or any fi, i<x, is final in q, then q ∉ HdL.Ffx[K].

      HdL.Ffx filters out all candidates in which fx or any lower head is final.
- Head fx is *final* in optima in Γ if ∀λ(Γ), CompL ≫ HdL.Ffx.
    - ∀q∈K, if any f is initial in q, then q ∉ CompL[K].

      CompL filters out all candidates in which any head is initial

a. Implication 1: If head fx is final then all lower heads are final.

- wsPA: fx is final in Γ if Px.α (= CompL ≫ HdL.Ffx) ∈ Γ. ∀i<x, Px.α => Pi.α by L-retraction: since Pi κα.dom ⊂ Px κα.dom, L(Pi.α) ⊂ L(Px.α) and W(Pi.α) = W(Px.α). Therefore, if Px.α is satisfied in Γ, then so is Pi.α, and fi is final in optima.
- nsPA: fx is final in Γ if Px.α (= CompL ≫ HdL.Ffx) ∈ Γ. If Px.α ∈ Γ, then by scope, ∀Pi, i<x, Pi.α (= CompL ≫ HdL.Ffi) ∈ Γ. If Pi.α ∈ Γ then fi is final in all optima.

b. Implication 2: If head fx is initial then all higher heads are initial.

- wsPA: fx is initial in Γ if Px.β (= HdL.Ff0 | …| HdL.Ffx ≫ CompL) ∈ Γ. ∀i>x, Px.β => Pi.β by W-extension: since Px κα.dom ⊂ Pi κα.dom, W(Px.α) ⊂ W(Pi.α) and L(Pi.α) = L(Px.α). Therefore, if Px.β is satisfied in Γ, then so is Pi.β, and fi is initial in optima.

- nsPA: fx is initial in Γ if Px.β (= HdL.Ffx $\gg$ CompL) $\in \Gamma$ or if $\Gamma \notin \Sigma$(Px). If $\Gamma \notin \Sigma$(Px), then P(x-1).β (=HdL.Ff(x-1) $\gg$ CompL) $\in \Gamma$, so f(x-1) is initial. By stringency, HdL.Ff(x-1)[K] $\subseteq$ HdL.Ffx[K]. Thus if HdL.Ff(x-1) is satisfied, so is HdL.Ffx, and fx is initial. Similarly, if Px.$\beta \in \Gamma$, $\forall$HdL.Ffi, i>x, HdL.Ffx[K] $\subseteq$ HdL.Ffi[K], so if HdL.Ffx is satisfied, then so is HdL.Ffi and all fi are initial in optima.

The logical derivation of the FOFC carries over to the following variations, which are intensional equivalents. While GEN and CON differ, the relationships between C filtration patterns are equivalent, and the resulting typologies the same.

4.3.2.2 AsymL

AsymL differs from SymL in both GEN and CON, but the logic of the argument and the intensional structure of the typology are equivalent. The system is summarized in the table below. In GEN$_{Asym}$ candidates all projection have the structure [$_{XP}$ (spec) [$_{X'}$ X comp]] where the head precedes the comp ('antisymmetric' syntax); head-final surface order results from comp-to-spec movement, as discussed above. The stringency scale is defined over the HdL Cs, but in this system, the antagonist is ObSp, which is satisfied when comp-to-spec movement fills a specifier.

22) *AsymL*

System	*SymL*	***AsymL***	*AsymO*
GEN	Sym	**Asym**	
EP scaled C	**HdL**		ObSp
Antagonist C	CompL	**ObSp**	HdL

AsymL is conceptually similar to BHR's analysis in adopts their anti-symmetric syntax analysis, with head-finality resulting from comp-to-spec movement. Their account drives movement with a feature, '^', similar to an EPP feature. In the present analysis, movement is driven by satisfaction of ObSp, which Grimshaw (2001:3-4) proposed to explain EPP.

In the VT in (23), lowercase grayed letters indicate lower (unpronounced) copies. As in SymL, all FOFC-violating candidates are HB (grayed, with bounders in the final column). Note that the HdL Cs assign more violations to some candidates here then to their surface-order equivalents in SymL, because of the additional structure of the copies; however, their filtration patterns remain the same, and the MOATs are isomorphic.

23) *AsymL VT*

Input	Output	HdL. Vf0	HdL. Vf1	HdL. Vf2	ObSp	HB-er
CP	a. [C [T [V O]]]				3	
	b. [C [T [O V o]]]	1			2	
	c. [C [[V O] T [v o]]]	1	1		3	b
	d. [[T [V O]] C [t [v o]]]	1	1	1	4	b (c)
	e. [C [[O V o] T [o v o]]	3	1		1	
	f. [[T [O V o]] C [t [o v o]]]	3	1	1	2	e
	g. [[[V O] T [v o]] C [[v o] t [v o]]]	3	3	1	4	e (f)
	h. [[[O V o] T [o v o]] C [[o v o] t [o v o]]]	7	3	1		

T_{AsymL} is extensionally surface-order equivalent to TSymL (assuming lower copies are not surface-apparent/pronounced): all and only the FOFC-satisfying candidates are possible optima in some language, as shown in the languages in (24). The Ts are intensionally equivalent up to C relabeling (bijection between CONS). The isomorphic Γs map to the same languages.

24) *Languages and Grammars of T_{AsymL}*

	Languages: optima	*Grammar (MIB)*	*# λ*
L1	Hd-initial: [C [T [V O]]]	`WeeL`	12
L2	V-final: [C [T [O V o]]]	`LWeL, LeeW`	4
L3	V- & T-final: [C [[O V o] T [o v o]]	`LLWL, LLeW`	2
L4	Hd-final: [[[O V o] T [o v o]] C [[o v o] t [o v o]]]	`LLLW`	6

Similarly, the PA is parallel, changing only the antagonist from CompL to ObSp. Both wsPA and nsPA are given in (25) and (26), respectively, with Ps named as above.

25) *wsPA(T_{AsymL})*

a. *Properties*

Property		α	β
P0	ObSp<>HdL.Vf0	LeeW	WeeL
P1	ObSp<>{HdL.Vf0, HdL.Vf1}.dom	LLeW	WWeL
P2	ObSp<>{HdL.Vf0, HdL.Vf1, HdL.Vf2.dom	LLLW	WWWL

b. *Value table*

	P0	P1	P2
L1	β	β	β
L2	α	β	β
L3	α	α	β
L4	α	α	α

26) *nsPA(T_{AsymL})*

a. *Properties*

Property		Scope	α	β
P0	ObSp<>HdL.Vf0		LeeW	WeeL
P1	ObSp<>HdL.Vf1	/P0α	eLeW	eWeL
P2	ObSp<>HdL.Vf2	/P1α	eeLW	eeWL

b. *Value table*

	P0	P1	P2
L1	β		
L2	α	β	
L3	α	α	β
L4	α	α	α

FOFC is derived under this analysis as in SymL, but differing in the antagonist driving head-finality. In AsymL, it is ObSp, satisfied by candidates where complements move to specifier positions in the projections, resulting in comp-head surface order. (As noted in fn4, CompL is not a possible antagonist in this system; Grimshaw 2001:24 (36) perspicuously shows that comp-to-spec movement increases violations of CompL, both doubling comp-internal violations and adding an intervening projection (in spec) for calculation of alignment violations.)

4.3.2.3 AsymO

AsymO is an inversion of AsymL: the two share GEN$_{Asym}$ and the two types of structural Cs in CON, but swaps their roles in the analysis. In AsymL, the stringency scale is defined over a set of HdL Cs with a single ObSp is the antagonist; in AsymO, a single HdL is antagonized with a set of stringently related ObSp Cs. The scale is built in the reverse order, with the least stringent isolating the *lowest* (lexical) head in an EP.

27) *AsymO*

System	*SymL*	*AsymL*	***AsymO***
GEN	Sym	**Asym**	
EP scaled C	HdL		**ObSp**
Antagonist C	CompL	ObSp	**HdL**

The VT is shown below, using the same notations as above.

28) *AsymO VT*

Input	Output	HdL	ObSp. Vf0	ObSp. Vf1	ObSp. Vf2	HB-er
CP	a. [C [T [V O]]]		1	2	3	
	b. [C [T [O V o]]]	1		1	2	
	c. [C [[V O] T [v o]]]	1	2	2	3	a, b
	d. [[T [V O]] C [t [v o]]]	1	2	4	4	a,b,(c)
	e. [C [[O V o] T [o v o]]	3			1	
	f. [[T [O V o]] C [t [o v o]]]	3		2	2	b,e
	g. [[[V O] T [v o]] C [[v o] t [v o]]]	3	4	4	4	b,e,(f)
	h. [[[O V o] T [o v o]] C [[o v o] t [o v o]]]	7				

T_{AsymO} is extensionally and intensionally equivalent to T_{AsymL} (with a bijection between the CONs), but with the inverse mapping between extensional languages and intensional Γs. In AsymL, L1 = {WeeL} (most stringent dominates antagonist: HdL.Vf0 ≫ ObSp) correlates with total head-*initiality* in optima. In AsymO, the extensional correlate is L1 = {WLLL} (antagonist dominates all: HdL ≫ OS.Vf0, OS.Vf1, OS.Vf2) but the intensional correlate is L4 = {LeeW} (most stringent dominates antagonist: ObSp.Vf2 ≫ HdL) with head-*final* order in all projections in optima. AsymO languages and Γs are shown in (29).

29) *Languages and Grammars of T_{AsymO}*

	Languages: optima	*Grammar (MIB)*	*# λ*
L1	Hd-initial: [C [T [V O]]]	WLLL	6
L2	V-final: [C [T [O V o]]]	LWee, WeLL	2
L3	V- & T-final: [C [[O V o] T [o v o]]	LeWe, WeeL	4
L4	Hd-final: [[[O V o] T [o v o]] C [[o v o] t [o v o]]]	LeeW	12

PA(T_{AsymO}) is likewise an inversion of PA(T_{AsymL}): each P*x*.α correlates with head-*initiality* rather than finality in f*x* and all higher projections in optima. Both wsPA(T_{AsymO}) and nsPA(T_{AsymO}) are shown in (30); the nsPA(T_{AsymL}) value table is repeated for comparison.

30) *PA(T_{AsymO})*

a. *wsPA(T_{AsymO})*

Properties

Property		α	β
P2	HdL<>ObSp.Vf2	WeeL	LeeW
P1	HdL<>{ObSp.Vf2, ObSp.Vf1}.dom	WeLL	LeWW
P0	HdL<>{ObSp.Vf2, ObSp.Vf1, ObSp.Vf0}.dom	WLLL	LWWW

Value table

	P2	P1	P0
L1	α	α	α
L2	α	α	β
L3	α	β	β
L4	β	β	β

b. *nsPA(T_{AsymO})*

Properties

Property		Scope	α	β
P2	HdL<>ObSp.Vf2		WeeL	LeeW
P1	HdL<>ObSp.Vf1	/P2α	WeLe	LeWe
P0	HdL<>ObSp.Vf0	/P1α	WLee	LWee

Value tables

AsymO	P2	P1	P0
L1	α	α	α
L2	α	α	β
L3	α	β	
L4	β		

AsymL	P0	P1	P2
L1	β		
L2	α	β	
L3	α	α	β
L4	α	α	α

The reason for head-finality in AsymO again rests on ObSp. Head-initiality is driven by satisfaction of the general HdL. The switch of the scale from one constraint type to another changes the subsets of heads in an EP referred to: for the HdL scale, sets are built top-down: a head f*x* is only included in a C*x* if all heads with *higher* f-values are. In the

ObSp scale, sets are built bottom-up: f*x* is only included in a C*x* if all heads with *lower* f-values are.

4.3.3 Typological equivalence and structural sensitivity to EP

The stability of the FOFC result under the variations above delineates the central components necessary for a theory of Con to entail the generalization. The core piece is a set of stringently-related structural Cs indexed to head positions along an EP. Within these parameters variation is possible, such as whether the stringency scale is defined on alignment (HdL) or obligatory element (ObSp) Cs. In all analyses, the general logic of stringency systems entails the FOFC results using recognized tools in OT analysis rather than stipulating a *FOFC constraint.[11] Additionally, more nuanced variations that still lack the stringency relation fail to derive the typology (see Appendix, §A.1, for alternatives).

The three systems instantiate variations of intensional and extensional typological equivalence. All have isomorphic MOATs and produce the same surface extensional languages. However, SymL differs from both Asyms in that the structures lacks copies. SymL and AsymL map the same Γ to the same (surface) language, swapping CompL and ObSp. AysmL and AsymO have exactly equivalent languages and Γs, but inverse mappings between these.

Grimshaw (2001 et seq.) has shown that the interactions of the structural Cs derive word order typologies, as well as economy of structure and movement effects. These

[11]Deriving FOFC with such a C is not straightforward. For a system with Con = {*FOFC, HdL, CompL} where *FOFC: *(x, y): [[x zp] y], *FOFC is satisfied by all candidates with FOFC-compliant orders; HdL and CompL are only satisfied by uniform orders—all initial or all final. All others candidates are HB, and T is defined by HdL <> CompL. *FOFC is not crucially ranked, unviolated in both Γs. This holds if *FOFC is violated by any pair of offending heads in the EP, not just successive pairs, since non-uniform orders are HB with CompL and HdL alone.

general Cs assess configurations in any projection, regardless of the head identity. These alone cannot generate FOFC because only candidates with uniform orders, either all head-initial or -final, satisfy them. Generating distinct orderings in distinct projections—such as the FOFC-satisfying non-uniform orders—requires targeted versions. Previous work proposing sets of targeted structural constraints includes Grimshaw's (2006) location-specific Cs picking out projections at the edges of matrix and subordinate clauses, and Steddy & Samek-Lodovic's (2009) projection-specific alignment Cs deriving Cinque's (2005) typology of DP-internal orders. Defining a subset scale over a sequence of heads as done here entails that while orders may differ in distinct projections, variation is contingent on the order realized by an immediately adjacent projection.

4.4 *BHR's analysis of FOFC*

BHR develop an analysis of the FOFC in the Minimalist Program (Chomsky 1995), and the typology was later developed in the Parameter Hierarchy theory (Biberauer et al. 2014, Biberauer & Roberts 2013, 2015, Biberauer & Sheehan 2012). The core component is a movement-triggering feature that results in head-final structures. Languages differ in whether this feature is present on a lexical head at the base of an EP and the degree to which it is inherited upwards by higher heads in the EP. As presented in the paper (p. 215 (77)), the analysis has four central components:

1) An antisymmetric analysis of word-final orders (Kayne 1994): all projections are underlying right-branching, $[_{XP}$ (spec) $[_{X'}$ X comp]], and head-finality results from comp-to-spec movement, assumptions followed here in $\textsc{Gen}_{Asym}$.[12]

[12]Antisymmetry is a crucial component of their analysis; the existence of SymL shows it is not critical to deriving FOFC in an OT system.

2) A strong locality condition on selection (Relativized Minimality, Rizzi 2001): feature inheritance is limited to the immediately-selected head. The movement feature can only be inherited by successive spans of heads in an EP and cannot skip any. (In the present analyses, non-skipping follows directly as a consequence of stringency systems. Candidates with a final-initial-final order sequence are not possible optima as a result of C interaction.)

3) The theory of Extended Projection (Grimshaw 2005): an EP defines the domain of the generalization. Selection depends on order in the EP.

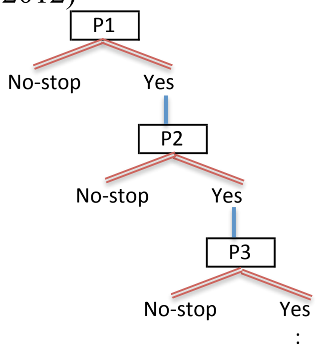
4) A general movement-triggering feature, represented by the diacritic ^: ^ triggers distinct types of movement when it occurs in combination with different sets of other features (p. 210). With an EP categorial feature F on the lexical head at the EP base, notated [F^], ^ triggers movement of the complement of that head to its specifier, resulting in head-final order.

The ^ feature can be inherited upwards, by each 'selecting' head. But as it can only be inherited with [F], its spread limited to spans of heads within the *same* EP. If the selecting head belongs to a distinct EP, it does not inherit [F] and consequently cannot inherit ^. However, [F] can be inherited independently of ^, so that a selecting head within the same EP may lack ^, while the selected head has it (deriving initial-over-final order). In this way, ^ can spread progressively upwards from the EP base, but once halted—not inherited—no higher head can inherit it. Head-finality occurs in the continuous span of projections whose heads have [F^].

This analysis defines in two dimensions of variation: a) the presence or absence of ^ on a lexical head L at the EP_F base, [F^] or [F]; and b) the extent to which ^ spreads up the EP if L has [F^] (the identity of the highest head inheriting ^) (p. 211). The first is a 'macroparameter' in that is has categorical effects in a language: absence of ^ entails total head-initial order in the language. The second corresponds to a set of parameters governing ^-inheritance for increasingly smaller subsets of heads in EP. These parameters are dependent on a [^F] setting of the first macroparameter; ^ can only spread in an EP if it is present on the base.

The Parameter Hierarchy structure of the FOFC typology (Biberauer et al. 2014, Biberauer & Roberts 2013, 2015), is organized it into a set of parameters governing the degree of head-finality in a language. Biberauer & Sheehan (2012:215) give a representation closely tied to the analysis by defining a set of parameters governing the presence or absence of ^ on heads in an EP, ordered from the lowest up. Their hierarchy is reproduced below, slightly modified by expanding the node they collapse with a recursive arrow and minor relabeling to facilitate comparison with the analyses here. P1 is a macroparameter that determines presence (yes) or absence (no) of ^ on lexical head L. Each subsequent P*x* determines whether the next higher selecting head $H_{x\text{-}1}$ has ^. Choice on P*x* is dependent on choice at P(*x*-1), because $H_{x\text{-}1}$ can only inherit ^ from $H_{x\text{-}2}$ if $H_{x\text{-}2}$ has it, corresponding to a 'yes' setting for P(*x*-1). Choice of a 'no' setting at any level stops inheritance of ^, and all lower parameters are 'moot'.

31) *FOFC word-order Parameter hierarchy* (adapted from Biberauer & Sheehan 2012)



Parameter 1 (P1): ^ is present on L
(L = lexical head at EP base)

Parameter 2 (P2): ^ is present on H_1
(H_1 selects L^)

Parameter 3 (P3): ^ is present on H_2
(H_2 selects H_1^)

…

This structure closely parallels that of the nsPAs treeoids, but departs from the generalized PH structure posited by ReCoS and shown in other publications (i.e., Biberauer et al. 2014:11), which use different parameters and ordering. The following section discusses this general structure and compares PH and PT.

4.5 *Property Theory and Parameter Hierarchies*

The analyses of FOFC in BHR and this chapter are both embedded in theories of the structure of linguistic typologies. Parameter Hierarchies and Property Theory have the shared goal of explicating the shape of the space of linguistic variation and propose the central organizing factor to be a set of formal binary choices, correlating with extensional traits. Order and dependencies between choices further structures the space, restricting possible combinations.

In terms of the FOFC analyses, both the Parameter Hierarchies analysis of BHR and the present proposal use a stringency-like element to derive the sensitivity of possible word orders to EP position. The OT analysis uses a set of stringency constraints. While stringency is not explicitly referenced in Parameter Hierarchies work, it is an inherent feature of the hierarchies, following from the way in which parameters determine the presence/absence of a feature over increasingly smaller subsets of heads.

Despite commonalities, the two theories come apart in significant ways. Under PH, there is a set parameter form and ordering in the hierarchy. The theory is separate from the details of a particular analysis, such as FOFC. In contrast, in PT, properties and their interdependencies result from the core pieces of the OT analysis: the constraints and their interactions. The theory brings out a structure that is emergent in all OT typologies, revealing their formal similarities and differences.

4.5.1 Parameter hierarchies

The theory of Parameter Hierarchies proposes a common syntactic typological structure[13], supporting it with analyses of a set of five cross-linguistic generalizations, among which FOFC figures prominently (Roberts 2010, 2012, Biberauer & Roberts 2015, a.o.). The empirical and theoretical work of ReCoS is a major contribution to typological study. The the project aims to "organise the parameters of Universal Grammar (UG) into hierarchies, which define the ways in which properties of individually variant categories may act in concert; this creates macroparametric effects from the combined action of many microparameters. The highest position in a hierarchy defines a macroparameter, a major typological property, lower positions define successively more local properties" (Roberts 2010:1). Typological properties arise from the combinations of the parameters, restricted by hierarchical ordering to rule out

[13] Roberts (2010:4) suggests that this structure is specific to *syntactic* typologies and that PF parameters concerning phonology and morphology are 'symmetrical', allowing for full logical combination of their settings. In contrast, PT finds core similarities between systems of diverse phenomena—such as phonology and syntax—but also differences between the structures of distinct syntactic systems, for example. This follows from the fact that PT does not specify a predefined structure, but explicates the structure within a typology.

unattested parametric options, gaps, that would be predicted from free cross-combination (Roberts 2015).[14]

Parameters have a common form: they govern the presence or absence of a feature F on some set of heads in the language. Parameter types differ in the particular set of heads assessed: a macro-parameter refers to all heads, microparameters to a natural-class defined subset, and meso- and nano-parameters to yet smaller subsets (Biberauer et al. 2014 (9)). Typological variation is defined in terms of which sets of features occur on which sets of heads.

Ordering between parameters follows a generalized uniform-branching binary tree structure. The nodes are labeled for the parameters, branching into yes/no choices of the setting ((32), Biberauer & Roberts 2013:22). One choice is decisive, the other leads to choice on a lower parameter.

32) *Generalized Parameter Hierarchy structure*

```
     F present?
No        Yes: present on all heads?
(none)
            Yes      No: which subset of heads
            (all)          …(some)
```

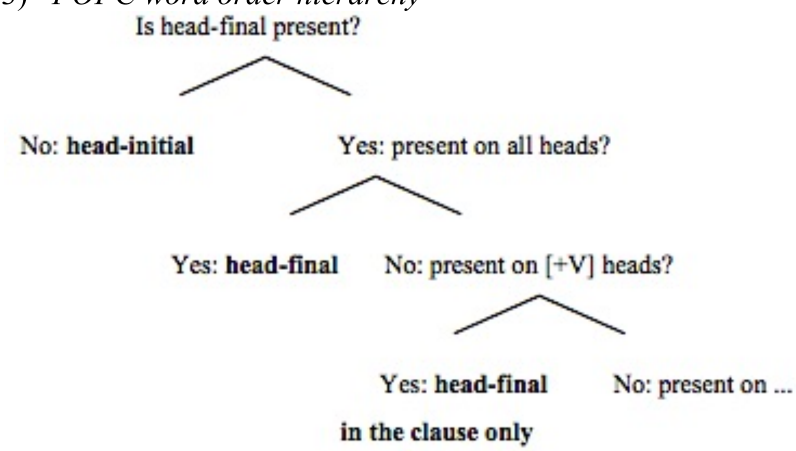
Higher nodes define the macroparametric options that have categorical effects on structures in a language. A language realizing one of the choices in the first two tree leaves has feature F on *no* or *all* heads. Choice on lower nodes depends on that at higher nodes; these are successively smaller parameter types determining F presence over increasingly smaller subsets of heads (Biberauer et al. 2014:11-12). Languages with settings of these have the particular trait in *some* set of structures. The hierarchy

[14]Baker's (2001) parameter hierarchies share some ideas with ReCoS. Baker analyzes a set of (morpho-)syntactic linguistic properties into a hierarchy and discusses parameter ordering in the context of learning.

partitions the typology by the degree of the trait correlated with F presence—as does a stringency system PA.

Under this theory, the FOFC analysis hierarchy is rendered as in (33) (Biberauer et al. 2014:11). In this representation, the parameter nodes are labeled for the extensional choices resulting from the parameter settings rather than the settings themselves, obscuring the details of the analysis; it is more explicitly brought out by replacing 'head-final' with feature ^. Parameters govern the presence of ^ on: all heads (No/Yes), then a subset, beginning with the lexical-feature-defined set of [+V] heads (all those in an $EP_{[+V]}$), and continuing to smaller subsets.

33) *FOFC word order hierarchy*



This structure produces a *none-all-some* sequence of options. A *no* setting on the highest parameter results in *none* of the trait occurring in the language; the relevant F is entirely absent in the language (i.e., no head-finality in (33)). A *yes* on the first two parameters generates a language with the trait in *all* relevant structures; the feature is on all heads. A *yes* on the first parameter and *no* on the second produces languages with the trait in *some* structures, necessitating choices on lower parameters to determine the particular set of heads with F.

While the five hierarchies analyzed differ in particular parameters, they are argued to adhere to the same core structure. However, as Biberauer et al. (2014:27) note, some of them depart from the generalized form, particularly among lower nodes, and strict adherence to the *none-all-some* initial sequence requires a 'no-choice' (monovalent) parameter in one, where one setting is cross-linguistically unattested (allegedly for 'functional' reasons, p. 29). Moreover, Biberauer & Sheehan's representation of the FOFC analysis (31) uses distinct parameters and ordering, for a *none-some-all* sequence. While *no* on P1 still correlates with *no* head-finality, *all* head-finality occurs in languages with *yes* settings for all parameters, not just for P2. This order closely matches the PA treeoid, but departs from the hypothesized *none-all-some* sequence, a discrepancy unaddressed in the cited works. This underscores the fact that the parameter hierarchy structure is not entailed by the pieces of the analysis itself, but from a separate theory.

The parameter form and ordering is proposed to arise from the interaction of three factors: UG, Primary Linguistic Data (PLD), and third-factor "domain-general acquisition strategies", specifically Feature Economy (FE) and Input Generalization (IG) (Biberauer & Roberts 2016:143). By FE, any feature that is not "unambiguously expressed by the PLD" will not be postulated (p. 145). In learning the FOFC word orders, the learner first hypothesizes that ^ does not exist in their grammar (*minimizing* the number of features), aligning with the first leaf on the parameter hierarchy tree, *none*. By IG, when unambiguous evidence exists, the learner *maximizes* use of the feature by postulating its presence on all heads. If head-finality occurs in the PLD, the learner swings to the assumption that *all* heads have ^, the second leaf of the tree. If further PLD shows some head-initial structures, the learner arrives at the *some* choice on the hierarchy, restricts the subset of heads considered, and repeats the steps (p. 148).

4.5.2 Property Theory

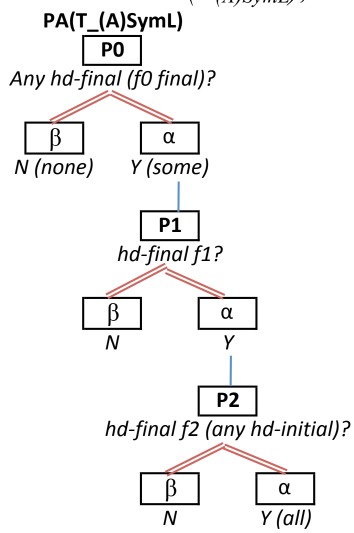
In PT, typological structure is decomposed into a set of properties whose values determine the intensional rankings in the Γs. The properties relate in system-specific ways, and are not guaranteed to conform to a specific structure. However, systems featuring stringently-related constraints share a common structure (ch 3), which classifies the typology into the same *none/all/some* sets of extensional choices as the Parameter Hierarchy structure. They are thus a basis for comparison, identifying theoretical parallels and differences.

Properties involving the most stringent constraint, C*n* (wsPs) correlate with macroparameters at the top of a hierarchy, whose settings/values have categorical effects across all relevant structures in the language. Properties involving less stringent Cs match lower parameter types; their values determine traits across smaller subsets of structures. In nsPAs, these Ps are hierarchically ordered, such that the (non)mootness of a nsP in a Γ depends on the value of a dominating P. The treeoid of nsPA($T_{(A)symL}$)[15] is repeated in (34), annotated for the correlated extensional choice following PH: each node queries head-finality of f*x* in f*x*P, from x = 0 to n (here, n = 2).

---

[15]PA(T_{SymL}) and PA(T_{AsymL}) treeoids are isomorphic. The treeoid in (34) differs from (18)c by listing the values in reverse lexicographic order. The parameter hierarchies could similarly be relabeled to use left-branching. While wsPAs define the same set of choices, they lack hierarchical organization; value combinations are limited by contradiction. Whether this alternative exists for Parameter hierarchies depends on defining parameter-setting contradiction. A flat structure also does not correlate with the desired hypothesized learning pathway.

34) *Treeoid nsPA($T_{(A)SymL}$), extensionally annotated*



A central difference between this representation and that of Parameter Hierarchy is the ordering of the properties/parameters. Following BHR (p. 172) in taking head-finality to be the 'marked' choice, parameter hierarchies (33) alternate between languages realizing extremes of a markedness scale: *no* head-final, then *all*, then *some*, repeating recursively in the *some*-set for subsets of heads. The parametric organization swings from least to most marked. This is based on a different markedness scale, where uniform order—all initial or final—is less marked than non-uniform, because it sets a macro- rather than micro-parameter (Biberauer et al. 2014:17). However, BHR's analysis defines presence of ^ as the marked case; thus the more heads bear ^, the more marked the language is predicted to be.

In contrast, in the PA structure, the languages are ordered from the least marked (*none* of the marked trait) to most (*all* of the marked trait), moving down. This structure is entailed by the stringency definitions of the constraints. The Ps cannot be reordered without other changes to the analysis itself; the order follows from the logical structure. In contrast, other orderings, including that of the PAs, are possible for the parameter hierarchy, as Biberauer & Sheehan's (2012) representation of the ^-based analysis (31) attests. The structure in (33) is not entailed by the analysis but by the general Parameter Hierarchy hypothesis, proposed to emerge from the interaction of UG, PLD, and general acquisition strategies. While the link between learning and typology is an important research area[16], this shows that hierarchies are imposed on a set of parameters, rather than arising from them.

In Property Theory, the structure emerges directly from the objects of OT itself: the interactions of CON over the space defined by GEN. The crucial constraint conflicts that define the grammars are properties, and their relations—the scopes—yield the hierarchical form. Rather than adhering to a predefined structure, the hierarchical relations represented in the treeoid are entailed by the typology and properties themselves, without appeal to outside learning factors other additional mechanisms.

4.6 *Predicting paths of diachronic change*

BHR show that FOFC constrains directionality of word order changes to certain pathways. No change can result in an FOFC-violating order. Thus "change from head-final to head-initial order in the clause must go 'top- down,' in that CP must be affected first, followed by TP, followed by VP [(35)a)]. Conversely, head-initial to head-final change must go 'bottom-up,' starting at VP before affecting TP, and then affecting TP before affecting CP [(35)b)]" (p. 192).

[16] Comparison of learning in the two theories is a topic of future work, using the learners of Tesar (2004, 2014) to learn the OT systems.

35) *Direction of diachronic change*

a. Head-final → head-initial

[[[OV] T] C] → [C [[O V] T]] → [C [T [O V]]] → [C [T [V O]]]

b. Head-initial → head-final

[C [T [V O]]] → [C [T [O V]]] → [C [[O V] T]] → [[[O V] T] C]

The directed step-wise change is exactly that predicted by the wsPAs by Alber's (2015ab) theory of diachronic variation as minimal property values change. Alber (2015a) develops the theory in an analysis of a stringency system sharing the structure of the systems developed here, so her analysis applies with little alteration. Two Γs are *property adjacent* (P-adjacent) if their descriptions in the PA differ in a single P value. Under the minimal change theory, change from one Γ directly to another is possible only if they are P-adjacent. When they are non-adjacent, differing in multiple values, change precedes stepwise via a path through other Γs, where each pair differs minimally. In this way, each step in the change path is a Γ, defined by a set of P values.

The present analysis predicts that only the pathways of change schematized in (35) are possible, as shown below using wsPA($T_{(A)symL}$) (36). In the sequences of value changes, a change from L1 to L2 (P0 value change) switches the ranking of CompL and the most stringent HdL.Ff0, resulting in head-finality in only the lowest projection, VP. Change from L1 to L4 must proceed through both L2 and L3, changing all P values one by one.

36) *Diachronic change as minimal P value change*

a. *(A)symL value table*

Γ: surface order	P0	P1	P2
L1: [C [T [V O]]]	β	β	β
L2: [C [T [O V]]	α	β	β
L3: [C [[O V] T]]	α	α	β
L4: [[[O V] T] C]	α	α	α

b. *Sequences*

Head-initial → head-final

L1 (βββ) → L2 (αββ) → L3 (ααβ) → L4 (ααα)

Change: P0 P1 P2

Head-final → head-initial

L4 (ααα) → L3 (ααβ) → L2 (αββ) → L1 (βββ)

Change: P2 P1 P0

The PA is crucial to predicting the change paths because it is this level of typological organization over which P-adjacency is defined. The result is not obtainable using *typohedral* (T) adjacency to condition minimal change, as all Γs are adjacent in this structure (see the typohedron in Appendix A.3 (41)). Using T-adjacency, change from any Γ to any other is predicted to be equally possible.

Defining P-adjacency and minimal change in the context of mootness (nsPAs) is more complex, as Alber (2015a,b) illustrates. Changing from a non-moot value to moot loses a P value; if retained, the Γ resulting from the change would be a refinement of the target Γ to which it was changing, with the additional value contributing an additional ranking. In the other direction, changing to a value requiring choice on a nsP results in adding a value; if the nsP value is not added, the resulting Γ is either a coarsening of the target Γ or not a Γ[17]. There is also the question of value choice: if either is possible, then a Γ with a great deal of mootness could change into several other Γs; in the nsPAs here, L1 could then change to any of L2, L3 or L4, failing to predict the paths BHR describe. See Alber (2015a,b) for further insight on mootness in this theory of variation.

4.7 *Summary*

The research program of the ReCoS project is a major step forward in typological analysis. The FOFC is significant both as an empirical discovery of possible cross-linguistically word orders, and as a target of theoretical explanation of linguistic typologies. This chapter proposed an analysis using a set of structural constraints defined in a stringency scale over an Extended Projection. FOFC follows from the logic of the systems realizing this scale. The typological structure emerges from property analysis. Predicted languages are defined extensionally by the degree of head-initiality/finality in syntactic phrases, aligning with intensional rankings characterized by the property values.

The analysis shares a central aim with that of BHR and the theory of Parameter Hierarchies. While employing different sets of tools and assumptions, both have a central stringency-esque core, where head-directionality in a given phrase is contingent on that of a higher or lower phrase. This is achieved in the present systems through the stringency scale, and in BHR's analysis through locality conditions on feature inheritance.

Both theories articulate the structure of the FOFC typology as a set of interdependent choices, parameters or properties, within broader theories of typological structure, Property Theory and Parameter hierarchies. Roberts (2013) argued that such parameter

---

[17]If, for example, the change resulted in a P value with a subordinated κ.sub; see ch 2.

hierarchies parallel OT typologies in limiting the space of variation.[18] However, the PH structure arises from a theory independent of, and additional to, the specific pieces of the analysis, appealing to external factors rather than being entailed by the parameters themselves. Property Theory explains the non-obvious but inherent typological structure of OT systems that emerges directly from the analysis.

A. *Appendices*

A.1. *Alternatives*

This appendix considers some alternative C sets that cannot produce the FOFC typology when used with the same the GEN. It does not constitute a categorical denial of the existence of alternative systems that depart more significantly from the assumptions here.

A.1.1. *Non-stringent head-specific Cs*

The first alternative defines HdL or ObSp Cs for *each* head in the EP individually, rather than using inclusion subsets. In the PAs, the general antagonist interacts with each specific C individually in a wsP (i.e., the PA structure resulting from making all Ps in nsPAs into wsPs). In AltSymL and AltAsymO, each of the eight candidates is possibly optimal (none HB) failing to derive the FOFC typology. The outlier is AltAsymL, which *does* generate the typology and thus would seem to refute the claim of the need for a stringency scale. However, because of roll-up movement and the fact that violations are assessed for all copies of a projection, a stringency relationship between the Cs is *derived* over the set of possible optima. The VT for AltAsymL is shown in (37).

---

[18]See Grimshaw's insightful response (2013b) to Roberts, esp. p. 577-8; this chapter follows the spirit of her critique, with reference to PT. See Grimshaw (2013a) for the analysis Roberts (2013) comments on.

37) *AltAsymL VT*

Input	Output	HdL.V	HdL.T	HdL.C	ObSp	HB-er
CP	a. [C [T [V O]]]				3	
	b. [C [T [O V o]]]	1			2	
	c. [C [[V O] T [v o]]]		1		3	a
	d. [[T [V O]] C [t [v o]]]			1	4	a
	e. [C [[O V o] T [o v o]]	2	1		1	
	f. [[T [O V o]] C [t [o v o]]]	2		1	2	b
	g. [[[V O] T [v o]] C [[v o] t [v o]]]		2	1	4	a,c,d
	h. [[[O V o] T [o v o]] C [[o v o] t [o v o]]]	4	2	1		

All candidates moving an f*x*P to spec f(*x*+1)P but *not* moving complement within the f*x*P (c, d, f, g) incur a violation of ObSp for each copy *and* a violation of the HdL C for the relevant head, f*x*. Possible optima are those where all moved complements have internal movement. This satisfies ObSp for f*x* and all lower projections, while deriving a stringency relationship between the HdL Cs: since neither f*x* nor any lower head is left-aligned, each HdL for a lower head is violated at least as much as the HdL for a higher head. Consequently, $T_{AltAsymL} = T_{AsymL}$.[19]

A.1.2. *HdL.EP*

Another alternative, from a suggestion from Grimshaw (p.c.), replaces the scale HdL Cs with a single C aligning all heads in an EP with the left edge of the entire EP (38). $\text{CON}_{HL.EP}$ includes a general HdL C, aligning each head in its own projection (this C is not crucial, as it and HdL.EP have the same filtrations). The higher the head in the EP, the more violations of HdL.EP incurred when that head is final. However, the system cannot generate the FOFC typology: only the two fully harmonic word order candidates, (a) and (h), are possible optima; no disharmonic orders are possible, FOFC-satisfying or not. The VT is shown in (39); all gray-shaded candidates are complexly HB by (a) and (h).

[19]This is not the case with AltAsymO: while movement doubles complement-internal violations, it satisfies the ObSp C specific to the projection to which it moved.

38) *HdL.EP*: *(X, f*x*): $[_{fnP}$ (…) X (…) f*x* where X = {f*i*, YP} (any head or projection)

Prose: for each pair (X, f*x*), where X is a maximal projection or another head, assign one violation if X intervenes between f*x* and the left edge of the EP, $[_{fnP}$.

39) *HdL.EP VT*

Input	Output	HdL	HdL.EP	CompL	HB-ers
CP	a. [C [T [V O]]]		3	3	
	b. [C [T [O V]]]	1	4	2	a&h
	c. [C [[V O] T]]	1	4	2	a&h
	d. [[T [V O]] C]	1	4	2	a&h
	e. [C [[O V] T]]	2	5	1	a&h
	f. [[T [O V]] C]	2	5	1	a&h
	g. [[[V O] T] C]	2	5	1	a&h
	h. [[[O V] T] C]	3	6		

While more nuanced than *FOFC, HdL.EP is similar in that it attempts to derive the condition through a single C rather than from the interaction of a set of Cs that together determine the order relations within an EP.

A.2. *Kiparsky 2015*

Kiparsky (2015)[20] develops an alternative OT analysis of FOFC word orders, in the context of a theory of syntactic change. His analysis uses a different set of Cs to derive the typology for a 3-hd EP, but similarities between the accounts further underscore the essential elements argued for here (see §4.3.3 above). Kiparsky's Cs are also in a stringency relationship; though there is no explicit reference to EP functional levels, the Cs derive a scale over three categories sequentially ordered in an EP: lexical head, intermediate functional heads, and the highest functional head (the complementizer). However, the equivalence between the analyses predictions come apart when the EP has more than 3 heads. Kiparsky's $\text{C{\small ON}}_K$ is below, as he states it.

[20]Thanks to Birgit Alber for reminding me of Kiparsky's analysis.

40) *Con$_K$*

Head-finality: heads follow their complements.

F≺XP: Functional heads precede their complements.

C≺XP: Complementizers precede their complements.

Harmony: If A is the complement of B, A and B have the same headedness.

Head-finality plays the same role that CompL does in SymL, satisfied by head-final projections. The two Cs F≺XP and C≺XP are correlates of HdL variants. The first refers to any functional head, thus including all EP heads except the lowest, lexical head. The second is specific to complementizers, (generally) the highest functional head. The Cs create a scale distinguishing three sections of the EP. This is a coarser scale than the f-value-based one used here. Because all non-complementizer functional heads are assessed equally, they must all be in the same order with respect to their complements in possible optima. The typology generates 4 Γs regardless of the number of heads in the EP input; when more than 3, the analysis cannot derive structures in which two such functional heads are in different orders (i.e., a T precedes its complement vP, but a v follows VP, [T[[VP]v]]). Kiparsky does not examine such structures.

The final C, Harmony, is violated by non-uniformity of head direction in the projection. It is crucial for the optimality of the all-initial candidate: since no C enforces hd-initial order of the lexical head, this structure is optimal only to satisfy the uniformity requirement when functional projections are hd-initial (satisfying the functional-specific Cs). No such C is needed in the present analyses, as the most stringent HdL C is violated by lexical head finality. Kiparsky's analysis, intended to explain historical change, motivates the lack of a similar C with the claim that "all languages are derived from a common OV proto-language" (2015:21). In a PA of Kiparsky's system, the property ranking Harmony and Head-finality is ns, with only Γs in which one of F≺XP or C≺XP dominates Head-finality (some hd-initial) having a value.

Kiparsky uses his analysis to explain diachronic word order changes using R-volume (Riggle 2010), a measure of Γ size as the number of λ(Γ)/CON!. In his theory, the most probable language is the one with the greatest R-volume; a learner is biased towards selecting the grammar consistent with previous evidence that has the highest R-volume. Full comparison between Kiparsky's and Alber's theories is beyond the scope of the present chapter.

A.3. *SymL: hedra and Γs*

The typology of SymL, T_{SymL}, is show on the 4C permutohedron in (41)a), mapping the constraints to {X, Y, Z, W}, and the Γs to the colors as indicated. The typohedron (flattened in (41)b) collapses all nodes (λs) within the same Γ to a single node (Merchant & Prince 2016). It is a tetrahedron, isomorphic to the typohedron of the 4C tops/bots (or T.1|*m*/T.*n*|1 in the terminology of DelBusso & Prince in prep.), though the systems are non-equivalent[21].

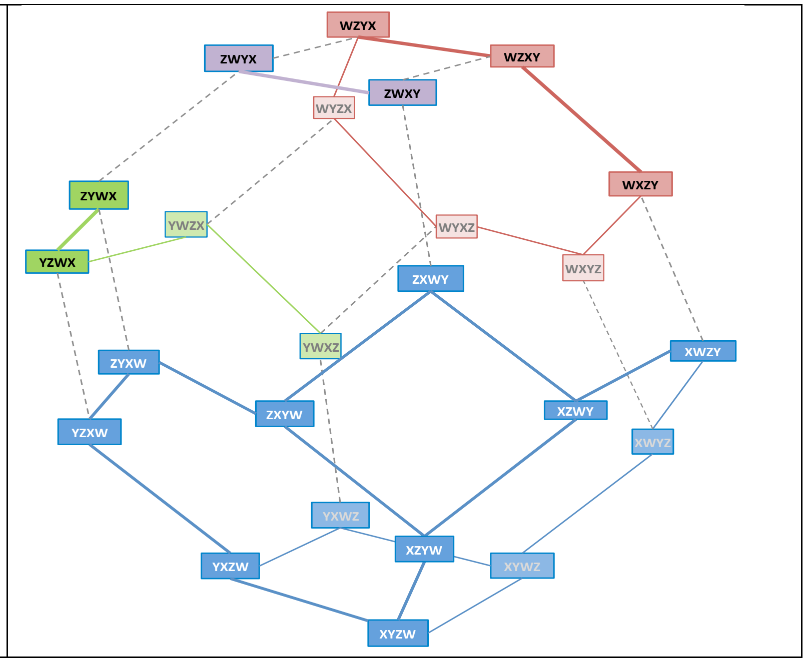
[21] See ch 3 on the typohedral isomorphism and permutohedral non-isomorphism of stringency Ts with *n*C scales and T.*n*|1/T.1|*m*.

41) *Hedra*

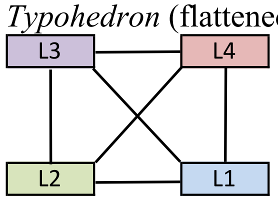
a. *Permutohedron*

Constraints:
X = HdL.Vf0
Y = HdL.Vf1
Z = HdL.Vf2
W = CompL

Grammars:
L1 = Blue
L2 = Green
L3 = Purple
L4 = Red

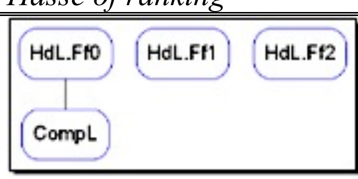
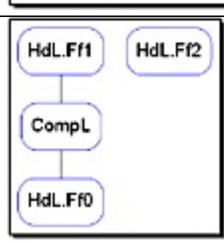


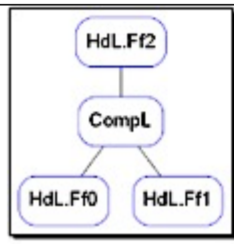
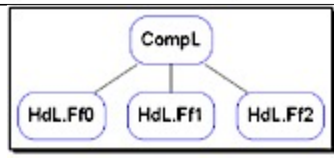
b. *Typohedron* (flattened, 2D)



Further details of the T_{SymL} Γs are given below; they differ in how many HdL scale Cs are dominated by CompL, corresponding to degree of head-finality in optima.

42) T_{Sym} *Γ*

L (# λ)	*Hasse of ranking*	*Notes*	*Extensional*
L1 (12)		HdL.Vf0 satisfaction (most stringent) => satisfaction of all less stringent HdL Cs (not crucially ranked)	No head-final projections (all initial).
L2 (4)		HdL.fV0 not satisfied; HdL.Vf1 dominates CompL; HdL.Vf2 is not crucially ranked.	Head-final in f0P (lexical base); all other initial.

L3 (2)		All HdL Cs are ordered relative to CompL; only the least stringent HdL.Vf2 dominates it.	Head-final in lower 2 projections (f0P, f1P).
L4 (6)		All HdL Cs dominated, unordered among each other; none satisfied in optima.	All head-final.

5 Conclusion

Formal typological analysis provides an otherwise unobtainable level of insight into both the theory and the linguistic data it seeks to explain. Well-defined generative linguistic theories predict typologies that seek to derive the extent and limitations on cross-linguistic variation in principled ways. The structure of the typologies shows how the theory explains the data.

This dissertation advanced the development of Property Theory (Alber & Prince 2016a, in prep.) and used it to understand core aspects of OT typologies. A factorial typology, often the final step in an analysis, is simply a list of languages, combinations of optima. Rather than an end, it is the starting point of analysis. In OT, such an unorganized list belies the fact that the typological space is highly structured, classifying sets of languages together and recognizing categories among them in systematic ways.

The formal results of the dissertation provide analytical tools that extend the reach and usability of property analysis. Addressing the question of the conditions under which a set of Ps yields a typology provides a way of assessing potential success of a given analysis, as well as diagnoses for failure (chapter 2). Examining the structure of a class of typologies sharing a common intensional structure rather than extensional topic shows that a broad range of systems explain diverse data in similar ways (chapter 3). This kind of analysis probes the formal objects of OT, specifically the MOAT, to identify key constraint relations that structure a typology. Chapter 3 proposed that these relations have both MOAT and property correlates. This opens the way for further developments to build a PA directly from a MOAT.

Property analysis further unifies traditionally separate subfields under a common theory of grammar: there is nothing inherently different about syntax and phonology in terms of this structure. The Final-Over-Final Condition (chapter 4) typology is explained by exactly the kind of stringency system studied in chapter 3. Its organization follows from the core constraint interactions, not from another hypothesized general form, as suggested by the theory of Parameter Hierarchies, though the theories recognize similar categories. Property Theory explains the emergent but non-obvious structure of OT typologies.

References

Alber, B. 2015a. Microvariation and typological properties. 1st Workshop on the Formal Structure of OT Typologies (WoFT 1), Rutgers University, May 29-30.

Alber, B. 2015b. Minimality and property analysis. Rutgers Optimality Research Group (RORG), Rutgers University, Oct. 7, 2015.

Alber, B. DelBusso, N. & Prince, A. 2016. From intensional properties to Universal Support. *Language: Phonological Analysis* 92.2: e88-116. ROA 1235.

Alber, B. & Prince, A. 2016a. Outline of Property Theory. Ms, University of Verona & Rutgers University.

Alber, B. & Prince, A. 2016b. Symmetries and dependencies in OT typologies. Workshop on Formal Typologies 2 (WoFT2), University of Verona, June 24-5, 2016.

Alber, B. & Prince, A. 2017. The book of nGX. ROA 1312.

Alber, B. & Prince, A. In prep. *Typologies*. Ms, University of Verona & Rutgers University.

Baker, M. 2001. *The atoms of language*. New York: Basic Books.

Bennett, Wm. G. 2017. Pulmonic venting and the typology of click nasals. Ms, University of Calgary. ROA 1308.

Bennett, Wm. G. & DelBusso, N. 2017. Typological consequences of ABCD constraint forms. In C. Hammerly, A. Lamont, B. Prickett, & K.A. Teztloff (eds.), *NELS 47: Proceedings of the 47th Annual Meeting of the North East Linguistic Society*, 1:75-88. Amherst, MA: UMass GLSA. ROA 1307.

Bennett, Wm. G. & DelBusso, N. To appear. Typological effects of ABC constraint definitions. *Phonology* 35.1 (2018). ROA 1326.

Bennett, Wm. G & DelBusso, N. In prep. ABC and ABP. Ms, Calgary University & Rutgers University.

Bennett, Wm. G., DelBusso, N. & Iacoponi, L. 2016. Formally mapping the typologies of interacting ABCD systems. In G. Hansson, A. Farris-Trimble, K. McMullin, D. Pulleyblank (eds.), *Supplemental Proceedings of the 2015 Annual Meeting on Phonology (AMP 2015)*. Washington, DC: LSA. ROA 1282.

Biberauer, T. 2017. The Final-over-Final Condition and particles. In M. Sheehan, T. Biberauer, A. Holmberg & I. Roberts (eds.) *The Final-over-Final Condition: A Syntactic universal*. Cambridge, MA: MIT Press.

Biberauer, T., Holmberg, A. & Roberts, I. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45.2: 169-225.

Biberauer, T., Holmberg, A., Roberts, I. & Sheehan, M. 2014. Complexity in comparative syntax: The view from modern parametric theory. In F. Newmeyer & L. Preston (eds.) *Measuring Linguistic Complexity,* 103-27. Oxford: Oxford University Press.

Biberauer, T. & Roberts, I. 2013. Emergent parameters: hierarchies and phylogenies. Advances in Phylogenetic Linguistics Conference. Ragusa Ibla, Italy, July 15-17, 2013.

Biberauer, T. & Roberts, I. 2015. Rethinking formal hierarchies: A proposed unification. *Cambridge Occasional Papers in Linguistics (COPiL)* 7: 1-31.

Biberauer, T. & Roberts, I. 2016. Parameter setting. In A. Ledgeway & I. Roberts (eds.) *The Cambridge Handbook of Historical Syntax*, 134-62. Cambridge: Cambridge University Press.

Biberauer, T. & Sheehan, M. 2012. Disharmony, antisymmetry, and the final-over-final constraint. In M. Uribe-Etxebarria & V. Valmala (eds.) *Ways of structure building*, 206-44. Oxford: Oxford University Press.

Brasoveanu, A. & Prince, A. 2011. Ranking and necessity: the Fusional Reduction Algorithm. *Natural Language & Linguistic Theory* 29.1: 3-70. ROA 794.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Cinque, G. 2005. Deriving Greenberg's Universal 20 and its exceptions. *Linguistic Inquiry* 36.3: 315-32.

Danis, N. 2014. Deriving interactions of complex stops. Ms, Rutgers University. ROA 1220.

DelBusso, N. 2015. Three basic typologies of reduplication. 1st Workshop on the Formal Structure of OT Typologies (WoFT1), Rutgers University, May 29-30.

DelBusso, N. 2016. Deriving reduplicants in Prosodic-Morphology Typologies. In A. Albright & M. A. Fuller (eds.), *Supplemental Proceedings of the 2014 Annual Meeting on Phonology (AMP 2014)*. Washington, DC: LSA. ROA 1251.

DelBusso, N. & Merchant, N. 2016. Aligning categories: Joins and properties. 2nd Workshop on Formal Typologies (WoFT2), University of Verona, Italy, June 23-24, 2016.

DelBusso, N. & Merchant, N. In prep. The join in Property Analysis. Ms, Rutgers University & Eckerd Colllege.

DelBusso, N. & Prince, A. In prep. The Weak Order Typologies (WOTs). Ms, Rutgers University.

DelBusso, N. & Prince, A. 2015. Syllable structure typologies. Ms, Rutgers University.

Erlewine, M. 2017. Low sentence-final particles in Mandarin Chinese and the Final-over-Final Constraint. *Journal of East Asian Linguistics* 26: 37-75.

Grimshaw, J. 2001. Economy of structure in OT. Tech. Rep., Rutgers University. ROA 434.

Grimshaw, J. 2005. *Words and structure*. Chicago: CSLI Publications.

Grimshaw, J. 2006. Location specific constraints in matrix and subordinate clauses. Tech. Rep., Rutgers University. ROA 857.

Grimshaw, J. 2013a. The structure of syntactic typologies. *Mind & Language* 28.4: 538-59.

Grimshaw, J. 2013b. Response to Roberts. *Mind & Language* 28.4: 573-8.

Hyde, B. 2012. Alignment constraints. *Natural Language & Linguistic Theory* 30:1-48.

Kayne, R. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kiparsky, P. 2015. Typologies as fitness landscape: Syntactic drift and convergence. 1st Workshop on Formal Typology (WoFT 1), Rutgers University, May 29-30, 2015.

Lombardi, L. 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *Natural Language & Linguistic Theory* 17: 267-302.

McManus, H. 2016. *Stress parallels in modern OT*. PhD dissertation, Rutgers University. ROA 1295.

Merchant, N. 2008. *Discovering underlying forms: Contrast pairs and ranking*. PhD dissertation, Rutgers University. ROA 964.

Merchant, N. 2011. Learning ranking information from unspecified overt forms using the join. *Proceedings of the Chicago Linguistics Society annual meeting 47 (CLS 47)*, 261-75. ROA 1146.

Merchant, N. & Prince, A. 2016/to appear. *The mother of all tableaux*. Advances in Optimality Theory. Sheffield & Bristol, CT: Equinox. ROA 1285 (citations to ROA version).

Merchant, N. & Riggle, J. 2016. OT grammars, beyond partial orders: ERC sets and Antimatroids. *Natural Language & Linguistic Theory* 34.1: 241-69. ROA 1158.

Mitchley, H. 2016. *Agreement and coordination in Xitsonga, Sesotho and Isixhosa: An Optimality Theoretic perspective*. Masters Thesis, Rhodes University.

Mitchley, H. & DelBusso, N. In prep. Typology of Agreement and Coordination in Bantu. Ms, Rutgers University, New Brunswick.

Prince, A. 2000. The special and the general. Talk given at the 24th Penn Linguistics Conference (PLC 24), University of Pennsylvania, Feb. 26, 2000.

Prince, A. 2001. Invariance under re-ranking. 20th West Coast Conference on Formal Linguistics (WCCFL 20), Feb. 23-25, 2001, University of Southern California.

Prince, A. 2002. Entailed ranking arguments. Tech. Rep., Rutgers University. ROA 500.

Prince, A. 2013. Voicing Typology. Class notes, Seminar in Phonology, Rutgers University.

Prince, A. 2016a. One tableau suffices. Tech. Rep., Rutgers University. ROA 1250.

Prince, A. 2016b. What is OT? Tech. Rep., Rutgers University. ROA 1271.

Prince, A. 2017. Representing OT grammars. Tech. Rep., Rutgers University. ROA 1309.

Prince, A., Merchant, N. & Tesar, B. 2007-2017. OTWorkplace. https://sites.google.com/site/otworkplace/.

Prince, A. & Smolensky, P. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell.

Prince, A. & Tesar, B. 2004. Learning phonotactic distributions. In R. Kager, J. Pater & W. Zonneveld (eds.), *Constraints in Phonological Acquisition*, 245-91. Cambridge: Cambridge University Press. (see also ROA 353).

Riggle, J. 2010. Sampling Rankings. Ms. University of Chicago. ROA 1075.

Rizzi, L. 2001. Relativized Minimality effects. In M. Baltin & C. Collins (eds.) *The handbook of contemporary syntactic theory*, 89-110. Oxford: Blackwell.

Roberts, I. 2010. Rethinking Comparative Syntax. European Research Council Grant #269752 proposal. http://recos-dtal.mml.cam.ac.uk/proposal/full-recos-proposal.

Roberts, I. 2012. Macroparameters and minimalism: A programme for comparative research. In C. Galves, S. Cyrino, R. Lopes, F. Sândalo & J. Avelar (eds.) *Parameter Theory and Linguistic Change*, 320-35. Oxford: Oxford University Press.

Roberts, I. 2013. Comments on Grimshaw. *Mind & Language* 28.4: 560-72.

Roberts, I. 2015. A parameter hierarchy for passives. Rutgers Linguistics Colloquium, Rutgers University, Nov. 6, 2015.

Samek-Lodovici, V. & Prince, A. 2002. Fundamental properties of Harmonic Bounding. Rutgers Center for Cognitive Science Tech. Rep. 71, updated as ROA 785.

Sheehan, M., Biberauer, T., Holmberg, A. & Roberts, I. 2017. *The Final-over-final condition*. Cambridge, MA: MIT Press.

Steddy, S. & Samek-Lodovici, V. 2011. On the ungrammaticality of remnant movement in the derivation of Greenberg's Universal 20. *Linguistic Inquiry*, 42.3: 445-69. ROA 1035.

Tesar, B. 1995. *Computational Optimality Theory.* PhD dissertation, University of Colorado Boulder. ROA 90.

Tesar, B. 2004. Using inconsistency detection to overcome structural ambiguity. *Linguistic Inquiry* 35.2: 219-53.

Tesar, B. 2014. *Output-driven phonology: Theory and learning*. Cambridge: Cambridge University Press.

Tessier, A.-M. & Jesney, K. 2014. Learning in Harmonic Serialism and the necessity of a richer base. *Phonology* 31.1: 155-78.