Introduction

Géraldine Legendre[a], Michael Putnam[b], Henriette de Swart[c], and Erin Zaroukian[d]

[a] Johns Hopkins University

[b] Pennsylvania State University

[c] Utrecht University

[d] École Normale Supérieure, Laboratoire de Sciences Cognitives et Psycholinguistique,

CNRS

The concept of *optimality* is a pervasive one in the study of the mind. Examples range from basic memory function (Anderson 1990), to computing the most likely 3D source for a 2D image in visual cognition (Geman and Geman 1984), to learning internal representations in neural networks (Rumelhart, Hinton, and Williams 1986), to Bayesian learning in any higher cognition domain (Oaksford and Chater 1998). These examples are ultimately all formalized in terms of competition plus an algorithm or procedure for choosing an output that optimizes a specified objective function. It is not surprising, then, that grammar (i.e., the native speaker's knowledge system) can also profitably be viewed as competition and optimization. An optimality-based approach also dovetails well with, on the one hand, the generative construal of grammar as a mental construct (I-language, Chomsky 1986) sharing fundamental architectural properties with other mental faculties, and, on the other hand, very basic empirical patterns that are observed cross-linguistically.

In this introduction, we start by sampling pervasive patterns in linguistic systems which result from optimality effects in syntax and semantics (Section 1.1). In Section 1.2, blocking, minimality, markedness, last resort, and complementarity are shown to be at the heart of an Optimality Theoretic analysis. The discussion also brings the directionality of grammar to the foreground: language generation is the process of selecting the optimal form for a given meaning (speaker perspective), and interpretation is the process of selecting the optimal interpretation for a given form (speaker perspective). The complex input-output relations in syntax and semantics raise the issue of how much speakers take the perspective of the hearer into account in language production, and how much hearers take the perspective of the speaker into account in language comprehension. In the last decade, exploration of this issue has resulted in movement away from architectures grounded in unidirectional optimization to some grounded in bidirectional optimization, as discussed in Section 1.3. This section includes a presentation of symmetric and asymmetric versions of bidirectional Optimality Theory in the syntax-semantics interface, as well as a discussion of the role of bidirectional optimization in first language acquisition. Section 1.4 offers an overview of the papers in this volume.

1.1 Optimality effects in natural language

Optimality effects − effects pointing to competition and optimization − cover a lot of ground in what is traditionally considered the purview of syntax. Accounts in either the (early) Minimalist Program (henceforth MP) or Optimality Theory (henceforth OT) have centered around the concept of *economy*. Empirical observations concerning blocking, last resort, and conflicts between principles or constraints, however, are not limited to

syntax (Section 1.1.1), but have come up in the semantics-pragmatics debate as well (Section 1.1.2).

1.1.1 Optimality effects in syntax

A wide range of syntactic phenomena involves competition and optimization processes, driven by principles of economy. Here we provide just a sampling of these phenomena under the headers of blocking, minimality, markedness, repair and last resort, complementarity, and conflicts between principles/constraints. These syntactic phenomena raise the following question: what architecture of the grammar is best equipped to coherently formalize the concepts of competition and optimization? These concepts are at the heart of these empirical patterns (Section 1.1.3), which prompts a discussion of OT as a suitable grammar architecture in Sections 1.2 and 1.3.

(a) *Blocking effects*: In pre-theoretical terms, blocking is observed when a default form fails to surface and is replaced by a marked form (cf. the Elsewhere Condition, Kiparsky 1973). While pervasive in (morpho)-phonology, blocking effects are also found in morpho-syntax and syntax more generally. To take one illustrative example of the latter, Hungarian has contrastive focus movement which in some cases results in verb particle stranding. Contrastively focused constituents move, say to [SpecFP] at the left periphery of the clause. In the presence of a moved contrastively focused constituent, tensed V moves to F, stranding the verb particle in situ (1). When the verb itself is focused, however, particle stranding is blocked and the PRT-V order is preserved (2).

(1)   a. *MARI hívta$_j$* [$_{VP}$ *fel t$_j$ Pétert*].

           Mary   rang   PRT Peter-ACC

         b. **MARI fel hívta Pétert.*

             Mary PRT rang Peter-ACC

           'MARY rang up Peter.'

(2)   *De,* [$_{VP}$ *én* [$_{VP}$ *ODA-VITTEM a levelet*]].

         But   I   PRT-took   the letter-ACC

         'But, I TOOK the letter THERE, (not brought it here.)'

The Hungarian stress rule is leftward-oriented. Szendroi (2003) argues that movement of

the focused constituent in (1) is triggered by the need to receive main stress. In (2) the

focused verbal complex stays in situ, because it is already in the main stress position, i.e.

aligned with the left edge of a phonological phrase. There is no trigger for its movement,

therefore, the PRT-V order is preserved. Szendroi's proposal relies on an implicit

comparison between two structures (+ V movement, − V movement) and economy.

Movement is grammatical only when necessary. Economy prevails.

(b) *Minimality effects*: One particular manifestation of economy, minimality/locality,

necessarily follows from a process of comparison and selection; it presupposes a set of

alternatives and a scale which orders them in terms of degree along a particular

dimension (e.g. distance). Since the earliest days of Generative Grammar (e.g.

Rosenbaum 1967; Ross 1967; Chomsky 1977) it has been known that locality effects are pervasive in syntax. We illustrate them here with the Superiority phenomenon in (3).

(3) a. I wonder [$_{CP}$ who$_1$ [$_{TP}$ t$_i$ bought what$_2$]].

b. *I wonder [$_{CP}$ what$_2$ [$_{TP}$ who$_1$ bought t$_2$]].

In the presence of two *wh*-phrases that may move to specCP in (3), the Superiority Condition (Chomsky 1973) states that only the *highest wh*-phrase (subject *who$_1$*) may move overtly. This condition, just like its early MP counterpart, Shortest Paths (Chomsky 1993), is a *transderivational* constraint, i.e. a constraint that relies on a comparison of alternative derivations, hence competition of sentences. The competition concerns which *wh*-phrase moves. The procedure for resolving it involves computing and comparing distance (see Chapter 1 of Müller and Sternefeld 2001 for further illustration and discussion). A related principle, Relativized Minimality (Rizzi 1990), which has dominated much discussion of *wh*-movement since the 1990's, also has the comparison built into the (set of) locality principles.

(c) *Repair and last resort effects*: To avoid 'crashing', or a failure to converge, language-specific operations typically do what must be done − no more and no less. The resulting structure may incur a cost, but this is worth paying because any alternative is worse. One example is *do*-support in English.

(4) a. John wrote books.

b. *John did write books.

c. John did not write books.

d. *John not wrote books.

To take one concrete economy-based analysis couched in MP, Bobaljik (1995) assumes that I contains an affix which morphologically combines with V under adjacency, yielding (4a). *Do* is not inserted in (4b) because it is not needed; economy prevails. If negation is present however, adjacency is blocked and *do* is inserted to support the affix (4c). By comparison, stranding an affix violates a basic principle of morphology and is worse than inserting *do* (4d). In an alternative OT analysis, Grimshaw (1997, 2013) argues that *do*-support results from the interaction of a (markedness) constraint favoring projections with heads outranking a (faithfulness) constraint disfavoring expletive elements.

(d) *Markedness effects*: A marked (or relatively uncommon) linguistic form is not ill-formed per se; it is marked only in comparison to other linguistic forms. For example, the unmarked value for syllable closure is open since all languages have open syllables (CV, V) while only a subset of languages allows closed syllables (CVC, VC). Similarly, all languages have subjects but only a subset of languages allows expletive subjects. In languages where these marked alternatives are possible, OT analyses suggest that there is a natural conflict between two types of constraints from which these patterns emerge (e.g. see Repair and Last Resort above).

(e) *Complementarity effects*: Complementarity is well-known to be at the core of the distribution of anaphoric elements (see also Section 1.3.1): the fundamental distinction is that anaphors (or reflexive/reciprocal pronouns) require local antecedents while pronominals (or non-reflexive pronouns) prohibit them, as shown in (5). Some limited overlap in distribution (6) and other cross-linguistic complications are real but the fundamental complementarity between anaphoric elements remains (Safir 2004).

(5)   a. The men$_i$ praised themselves$_{i/*j}$.

        b. The men$_i$ praised them$_{j/*i}$.

(6)   a. The men$_i$ bought their$_{i/j}$/each other$_i$'s pictures.

        b. These men$_i$ dislike pictures of themselves$_i$/them$_{i/j}$.

Building on Reinhart (1976), the original Binding Theory (Chomsky 1981) relied on an *identical* local binding domain within which an anaphor must be bound (Principle A) while a pronominal must be free (Principle B). In response to empirical challenges, alternative theories of sentential anaphora were developed whereby Principle B was eliminated as an independent principle. Ignoring details of implementation, this alternative approach gives center stage to the theoretical concept of competition, with Principle B effects explained as the natural outcome of being blocked by or losing the competition to an optimal anaphor (e.g. Hellan 1988; Burzio 1989, 1991; Safir 2004; Hendriks this volume; amongst others).

(f) *Effects of conflict between principles/constraints:* As discussed in Diesing (1996), sentences with object shift in Icelandic have different interpretations than their counterparts without object shift.

(7)   a. Í prófunum svarar hann sjaldan erfiðustu spurningunni

         in exams-the answers he rarely most-difficult question-the

       b. Í prófunum svarar hann erfiðustu spurningunni sjaldan   (OS/object shift)

         in exams-the answers he most-difficult question-the rarely

The (narrow scope) interpretation of (7a) is that regardless of which exam he is taking, he rarely answers whichever question happens to be the most difficult one in that particular exam. The (wide scope) interpretation of (7b) is that there is one particular question which is more difficult than all others and which appears in all exams, and when he encounters it, he rarely answers it. In contexts where object shift is not possible (e.g. compound tenses), the sentence actually allows both interpretations (Vikner 2001).

(8)   a. Í prófunum hefur hann sjaldan svarað   erfiðustu   spurningunni

         in exams-the   has he   rarely answered most-difficult question-the

       b. *Í prófunum hefur hann svarað erfiðustu   spurningunni sjaldan   (OS)

         in exams-the   has he answered most-difficult question-the   rarely

As Vikner (2001: 327) puts it, "what matters is not just whether the object has undergone object shift or not, but also whether it could have moved *if it had wanted to*". By economy (of movement/Procrastinate) a Diesing-type analysis (object within VP) entails

a possible narrow scope interpretation for both (7a) and (8a). A wide scope interpretation

in (7b), on the other hand, follows from movement outside of VP in violation of

economy. (8a) with the object still in VP can have a wide scope interpretation despite

violating Diesing's Scoping Condition because there is no alternative. Thus an account of

(7)-(8) together requires abandoning the view that the Scoping Condition is a condition

on Convergence, hence inviolable.

In sum, there is little doubt that optimality effects are pervasive in syntax and its

connection with interpretable interfaces. The real debate surrounds the implications of

these effects for the architecture of grammar. We will come back to the architectural

issues in Section 1.1.3, after reviewing similar optimality effects in semantics/pragmatics.

1.2 Optimality effects in semantics and pragmatics

The semantics and pragmatics literature provides numerous examples of optimality

effects similar to those discussed above in syntax. In fact, for any given linguistic

phenomenon, optimality effects often have two faces: a syntactic one (discussed in

Section 1.1.1) and a semantic/pragmatic one, which is the focus of this section.

(a) *Effects of conflict between principles/constraints*

The interpretation of anaphoric pronouns is sensitive to morpho-syntax (phi-features) and

accessibility constraints on the antecedents (see discussion of *Minimality effects* and

Complementarity effects in Section 1.1.1 above). Often, this leaves more than one

potential antecedent for the pronoun, at which point listeners use different *interpretive*

strategies to resolve the anaphoric relation. Building on a long-standing discussion in the

literature on pronoun resolution, Kehler and Rohde (2013) contrast the following examples, with preferred interpretations of pronouns shown in brackets:

(9)   Mitt narrowly defeated Rick, and the press promptly followed him to the next

        primary state.   [him = Mitt]

(10)   Mitt narrowly defeated Rick, and he quickly demanded a recount.   [he = Rick]

(9) illustrates a preference for subjects as the antecedent of a pronoun, as argued for by centering theory (Grosz, Joshi, and Weinstein 1995). In (10), however, this preference is overruled by world knowledge and abductive reasoning (Hobbs et al. 1993). Note that (9) is a narrative sequence, while (10) involves causality, so discourse structure is relevant. Kehler and Rohde (2013) develop a probabilistic model in which top-down expectations about coherence relations and what is coming up next are balanced with bottom-up linguistic evidence about topic-hood status of potential referents.

(b) *Complementarity effects*

In the literature on binding, anaphors and pronominals are taken to have a complementary syntactic distribution (see Section 1.1.1. above). Complementarity holds for their interpretation as well, as reflexives like *himself* are encoded as identical to another argument in the same clause, whereas pronouns like *him* cannot be:

(11)   a.   John$_i$ saw him$_{*i/j}$/himself$_{i/*j}$ in the mirror.

        b.   John$_i$ thinks everyone$_j$ likes him$_{i/*j}$/himself$_{*i/j}$.

Another example (less intertwined with syntax) pertains to the complementary interpretations of English definite and indefinite articles. Definite articles have a unique/familiar discourse referent, while indefinite articles have a non-unique/new discourse referent. For example, the indefinite article in (12i) introduces a new discourse referent:

(12)   i. A couple entered the church.

        ii. The man immediately sat down, but the woman looked around for a long time before she chose a seat.

In order to link the individual man and woman in (12ii) to the couple introduced in (12i), a definite article is appropriate, since an indefinite article would suggest that we are talking about unrelated people. In terms of Horn (1984), *<a, the>* define a Horn scale: the definite article has a *stronger* meaning than the indefinite article because only the definite article conveys uniqueness/familiarity. If the speaker uses the indefinite article, this conveys lack of uniqueness as the result of a scalar implicature. That is, the hearer reasons that the speaker did not choose the stronger definite article because uniqueness/familiarity does not hold in this case (Hawkins 1991). Heim (1991) formalizes the need to choose the strongest expression as the principle of Maximize Presupposition.

(c) *Markedness effects*

Most discussions of markedness deal with morpho-syntactic complexity or typological and frequency distributions (see Section 1.1.1, and Haspelmath 2006 for an overview). Economy considerations of length, ease of pronunciation, or frequency are oriented towards forms. Semantic/pragmatic markedness instead typically involves contrasts in meaning. In the domain of animate nouns we see pairs where one member is used for the male referent and the other for the female referent, e.g. *<actor, actress>, <hero, heroine>*. Going back to insights by Jakobson (1984), the female *actress/heroine* is often considered the marked member of the pair, because it is specified as female. The male *actor/hero* is the unmarked term, because it either takes on the general meaning (referring to male+female members of the category) or the male meaning, complementary to the female meaning. This type of semantic markedness is traditionally accounted for in terms of lexical features, where the specification of *actress* as female blocks the female meaning for *actor* (cf. Dalrymple and Kaplan 2000 for an analysis of grammatical gender in these terms).

Semantic markedness is not restricted to lexical semantics. Horn (1989/2001: 161 sqq) cites psycholinguistic evidence concerning negation in L1 acquisition and processing difficulties with negation as suggestive evidence in favor of the semantic markedness of negation. Negation is a universal category of natural language (Dahl 1979), but negative structures are syntactically more constrained than their affirmative counterparts (Gívon 1979). Markedness of meaning is thus typically related to markedness in form: all languages have a marker of negation, but few have a marker of affirmation. Blocking effects arise in the distribution of negative/affirmative meanings

over marked/unmarked forms: lack of marking leads to an affirmative, rather than a negative interpretation of the sentence.

(d) *Blocking effects*

Total blocking arises when some expression, despite having an optimal interpretation, does not exist because a preferable expression does (see Section 1.1.1). But sometimes blocking is just partial. According to McCawley (1978: 257), *kill* and *cause to die* have the same linguistically encoded semantic structure, yet they make different contributions to the discourse because of general principles of cooperative behavior (Grice 1975). The weakening and strengthening of meanings correlate with the well-known Q-principle ('Say as much as you can') and R-principle ('Say no more than you must') in pragmatics. Horn's (1984) division of pragmatic labor accounts for such blocking effects: the unmarked meaning of 'direct action' is expressed by the unmarked form *kill*, restricting the marked form *cause to die* to the marked meaning of 'indirect action'. More complex instances of partial blocking include the relation between lexical and syntactic negation in *unhappy/not unhappy*, where *not unhappy* indicates 'slightly happy' rather than plain 'happy'. This phenomenon, known as litotes, indicates that double negation in natural language has pragmatic effects not captured by the equivalence of ¬¬p and p in propositional logic.

(e) *Last resort and repair*

Chierchia (1998) introduces the operation of type shifting as a last resort in the debate on bare nominals. In a type-theoretical perspective on natural language, both common nouns

(N or NP) and verb phrases denote properties (type <e,t>), and as such the two cannot combine to yield a truth value (t). Articles serve as type shifters to enable nominals to serve in argument position. Articles denote functions of type <<e,t>,<e,t>,t> or <e,t>, e>. Thus they take N/NPs of type <e,t> and produce DPs of one of two types: 1) type <<e,t>,t> (generalized quantifiers), allowing the DP to take the VP it as an argument, or 2) type e (entities), allowing the VP to take the DP as an argument. This is evident in English, where articles like *the* and *a* are mandatory, and bare singular count nouns cannot appear in regular argument position (13a):

(13)   a.   *(A/The) coconut is on the table.

       b.   Coconuts are healthy.

In terms of Generalized Quantifier theory, the indefinite article in (13a) shifts the type <e,t> expression denoted by the noun to the type <<e,t>,t> denotation of a generalized quantifier DP, and assigns it an existential interpretation. The definite article in (13a) shifts the type <e,t> expression denoted by the noun to a type e denotation of an entity via the iota operator (ι). In order to account for bare nominals in regular argument position, such as the bare plural in (13b), Chierchia (1998) invokes type shifting as a last resort: for any type shift operator τ and any X, the free type shift τ(X) is blocked if there is an overt determiner D that encodes the meaning τ. Free type shift operations that are not lexically encoded are thus available as a last resort strategy that repair the type mismatch between the type <e,t> denotation of the N/NP and the type e or type <e,<e,t>> denotation required for the nominal in regular argument position.

In (13b), the bare plural gets a generic interpretation, which Chierchia takes to be kind referring, following Carlson (1980). Kinds are viewed as special entities in the ontology, so the bare plural denotes an expression of type e. Formally, Chierchia uses the 'down' operator $^\cap$, which maps properties (of type <e,t>) onto kinds (of type e) to interpret the bare plural in (13b) in terms of $^\cap$COCONUT. The down operator is not lexically encoded in English, so it is available as a free type shift. Given that $^\cap$ builds the maximal sum out of all entities realizing the kinds across worlds, it applies to plurals (13b), but cannot repair the singular bare noun in (13a).

In sum, optimality effects such as complementarity, markedness, blocking, and last resort are widely discussed in the semantics/pragmatics literature. Principles that have been proposed to account for these phenomena include probabilistic models of top-down and bottom-up interpretation for anaphora resolution (Kehler and Rohde 2013), Horn's division of pragmatic labor for the distribution of marked/unmarked meanings over marked/unmarked forms (Horn 1984), Heim's Maximize Presupposition for complementarity effects (Heim 1991), and Chierchia's type shifting as a last resort to repair type mismatches (Chierchia 1998). The pervasiveness of optimality effects in both syntax and semantics raises important questions about the architecture of grammar.

1.3 Implications of optimality effects for the architecture of grammar

Optimality effects in syntax and semantics/pragmatics raise the question of whether the theory of grammar should include extra principles of economy, complementarity, and competition *with the comparison built in* (e.g. Procrastinate, Relativized Minimality, Maximize Presupposition, Type-shifting as a last resort) or

whether such notions should be *theorems* that follow from completely general principles of a competition-based grammatical architecture. The Minimalist Program and LF-oriented semantic theories tend to take the former view while OT takes the latter. It is worth noting that building the comparison into the statement of principles or constraints leads to rather complex statements of these principles and constraints. Such complications are not obviously in line with parsimony, but they make it possible to maintain inviolability (aside from possible parametric variation), and compositional semantics can be read off (the LF level of) a syntactic tree. OT takes the view that optimality results instead from the resolution of conflicts amongst formally simple and violable constraints. The debate ultimately rests on whether competition is the norm (e.g. in the very architecture of OT) or the exception (e.g. in the specific principles of MP or LF semantic approaches that induce it).

A further fundamental architectural issue, mostly discussed in syntax, is whether transderivational/translocal constraints exist, which involve competition among derivations/representations, and the concomitant question of whether the grammar is derivational or representational. This issue does not distinguish MP or OT since there are representational and derivational versions of both, and some MP principles have been formulated from either perspective (e.g. Relativized Minimality vs. the Minimal Link Condition). As emphasized by Prince and Smolensky (1993/2004: 95-6), OT per se makes no commitment to a *serial* or *parallel* architecture. In fact a serial theory − Harmonic Serialism − was developed in the foundational work. For syntax more specifically, Legendre, Smolensky, and Wilson (1998: 285-6) unambiguously state that OT is quite compatible with a derivational theory and most analyses developed in one

perspective can be rephrased in the other (see Hale and Legendre 2004 for an example of recasting a derivational remnant movement analysis of German VP fronting into a representational one).

Both early OT (including OT-LFG) and early MP unsurprisingly tended to make use of global constraints as they grew out of representational theories with output constraints (GB, LFG). Chomsky (1991, 1993, 1995, 1998) initially proposed transderivational constraints in MP (e.g. Fewest Steps, Shortest Paths, Procrastinate, etc.) designed to choose the most economical derivation in the reference set. Close scrutiny of such constraints led Müller and Sternefeld (2001: 27) to conclude that "a minimalist syntax with [transderivational] constraints has exactly the shape that Prince and Smolensky (1993/2004) attribute to an optimality-theoretic grammar": it includes a generator governed by *local* constraints responsible for creating the reference (or candidate) set and an evaluator determining the optimal candidate or derivation.

The recent trend in MP has witnessed a shift away from global constraints. In some respects, this shift away from transderivational constraints have led to versions of the MP in which Merge, i.e., the operation responsible for building syntax structure, is free, resulting in outputs that must be evaluated for their well-formedness (see models such as Borer 2013 and Boeckx 2014 for example), signaling perhaps a return to a filtering architecture similar to GB-era. Within OT, Gereon Müller and colleagues have developed a derivational version − their paper in the present volume specifically argues for a crucial role of 'local' optimizations (see Müller 2011 for an exhaustive treatment of this topic). From a broader perspective, the bottom-up derivational approach resulting in right-to-left structure building − whether it integrates violable principles/constraints or

not − remains difficult to reconcile with a very basic cognitive reality − the real-time left-to-right processing of syntax (Phillips and Lewis 2013).

1.2 Optimality Theory as the basic architecture of the grammar

At the outset it is worth emphasizing that OT is *not* a substantive theory of (any component of) the grammar. It is a formal theory of constraint interaction *only*. OT is therefore fully compatible with any formalized theory of syntax, including constraint-based theories like LFG. The substantive *content* of principles/constraints ultimately derives from empirical generalizations gathered from at least 60 years of linguistic investigations. It is also worth noting that, like other linguistic frameworks, OT is a program of research with many theory-internal issues still unsettled. This section presents the basic principles of OT grammar (Section 1.2.1), and applies them to syntax (Section 1.2.2) and semantics/pragmatics (Section 1.2.3).

1.2.1 General features of an OT grammar

Optimality Theory rests on four basic assumptions:

(14)   a. *Universality*: Constraints are universal.

b. *Violability*: Constraints are violable.

c. *Ranking*: Constraints are ranked.

d. *Comparative evaluation*: Optimal linguistic expressions are determined by

comparing potential linguistic expressions or candidates for a given input

(candidate set).

Evaluation of candidates is based on strict domination among a ranked set of universal constraints, contrary to its predecessor, Harmonic Grammar, which exploits constraint weighting (Legendre, Miyata, and Smolensky 1990a, b; Smolensky and Legendre 2006; Pater in press). For any two constraints C_1 and C_2, either C_1 outranks C_2 or C_2 outranks C_1. Only the optimal candidate is grammatical.

Optimality is defined as follows: A is more harmonic than B iff A better-satisfies the highest-ranked constraint that distinguishes A from B. The procedure H-EVAL (Harmony Evaluation) which selects the most harmonic candidate (i.e. with the best violation profile) as the optimal candidate is in fact only one component of the grammar. This component is fed by GEN(ERATOR) which relies on inviolable constraints to generate the candidate set of competing outputs from an input which, theory-internal issues aside, resembles the numeration in MP. Broekhuis and Vogel (2013) argue that many proponents of MP and OT adopt more or less the same global architecture of the grammar represented in (10). In MP the generator is the computational system C_{HL} and the evaluator consists of interface conditions (Chomsky 2000 and later). In OT the evaluator is H-EVAL.

(15) The architecture of grammar (Broekhuis and Vogel 2013)

[INSERT EX 15 HERE]

Given universal constraints and language-particular ranking, a *factorial typology* is generated for the set of all possible rankings (possibly subject to some ranking restrictions hypothesized as part of UG).

Some confusion lingers as to the predicted set of grammars in OT, possibly because the effect of *harmonic bounding* on an OT candidate set is not sufficiently appreciated. While it is true that GEN takes an input and *in principle* could return a large candidate set, the actual comparative evaluation operates on non-harmonically bounded candidates only. Any candidate that incurs a superset of the violations incurred by a competitor is a *harmonically bounded* candidate (Prince and Smolensky 1993/2004; Samek-Lodovici 2001; Samek-Lodovici and Prince 2002). It can never be optimal under any ranking and is effectively not part of the competition.

As a simple illustration, having more structure (e.g. CP) may offset constraint violations (e.g. a constraint requiring *wh*-phrases to surface in specCP due to their operator status) but beyond trying to beat competitors with respect to satisfying a finite set of constraints, adding further structure (e.g. extra CPs) only results in worse competitors. Thus, for a set of N constraints, N! (N factorial) is never relevant to the number of languages predicted. There are N! rankings but not N! languages because (a) there are very few candidate outputs to choose from, due to harmonic bounding, and b) as a consequence many rankings produce the same optimal output.

OT relies on an input-output mapping architecture. In semantics/pragmatics the input is an expression (or form) and candidate outputs are alternative interpretations of this expression. The reverse holds in syntax: the input is a meaning, typically assumed to consist of predicate-argument structure, lexical items, information structure features, and

scope (at least in some proposals); candidate outputs in syntax are alternative expressions of this meaning.

OT relies on two types of constraints: *markedness* constraints, like well-formedness constraints in other frameworks, evaluate the inherent goodness of an output structure; *faithfulness* constraints regulate the input-output mapping and are unique to OT. Input-output faithfulness constraints limit how far candidate outputs may differ from the input. They require the output to express all and only the properties of the input. Faithfulness and markedness constraints are inherently conflicting. Preserving a contrast encoded in the input via a feature will typically lead to some cost in markedness.

Faithfulness constraints are crucial to the OT conception and have played a pivotal role since the theory's inception. In syntax their role is central to the treatment of ineffable meanings and absolute ungrammaticality (e.g., of multiple questions 'who ate what' in some languages; see Legendre et al. 1998; Legendre 2009; Rice and Blaho 2010 for further discussion). While ineffability has been construed by some as evidence for clash & crash syntax (Pesetsky 1997), it alternatively points to a genuine role for input-output faithfulness in syntax. The overall scope of faithfulness in syntax largely remains to be explored.

A foundational faithfulness constraint that plays an important role in OT semantics is F(ULL)-INT(ERPRETATION). F-INT is proposed by Zeevat (2001) as a semantic constraint which states that morphemes and lexical items have meanings, and the hearer has to faithfully interpret all the speaker says. The combination of the faithfulness constraint F-INT and the markedness constraint *INVENT (do not add material to the content of the sentence without a proper reason) restores important features of

compositional semantics. Zeevat (2001) considers these two constraints crucial

ingredients of a grammar in which OT syntax and semantics pair up to define the syntax-

semantics interface (more on bidirectional OT in Section 1.3). The widespread role of

faithfulness constraints in OT semantics may also be related to the observation that

semantic markedness seems harder to define than syntactic markedness (see Section 1.1.2

above).

While the optimization-based view on grammar originated in the domain of

phonology and syntax, it quickly spread to semantics and pragmatics. The next two

subsections focus on unidirectional approaches, that is, mapping meanings onto forms

(OT syntax, Section 1.2.2) or mapping forms onto meanings (OT semantics, Section

1.2.3).

1.2.2 OT syntax

This section highlights unique properties of OT syntax. Comparison with MP may be

made to clarify where the two approaches differ and what these differences entail. The

discussion eschews the issue of representations vs. derivations for the reasons stated in

Section 1.1.3 above, as well as theory-internal issues that are discussed in more

comprehensive presentations of OT syntax (e.g. Legendre 2001; Müller 2009). OT is

primarily motivated by the existence of cross-linguistic variation in syntax as a result of

its interfacing with other modules of the grammar and the view that variation arises as a

consequence of universal constraints conflicting with one another, these conflicts being

resolved by language-specific constraint ranking. In contrast, cross-linguistic variation in

MP is handled via featural properties of individual lexical or functional items (e.g.

presence/absence of EPP features) and divorced from interface conditions; variation is thus relatively unconstrained.

Consider the case of the syntax-prosody interface (and its explanatory power in accounts of word order variation) as an illustrative example. Linguistic expressions are generated by optimally efficient derivations that must satisfy conditions that hold at interface levels in MP. For information structure this means, for example, that focus and topic are instantiated as formal features on syntactic nodes, which interface with their phonological form via rules of information structure realization (prosody) and with their discourse status via rules of information structure interpretation (e.g. Büring 2007). In OT, focus and topic are features on particular elements in the input to optimization (e.g. Samek-Lodovici 2005, in press). Both analyses must contend with the fact that some movements are triggered by prosodic constraints. The main difference lies in the underlying architectural claims. In his MP analysis Büring is led to posit one of two theoretical complications: (a) either the syntax anticipates aspects of the prosodic structure (perhaps along the lines of Zubizarreta 1998) or (b) optional movements must be allowed that are filtered under a matching condition at the syntax-prosody interface. In Samek-Lodovici's OT analysis this is accomplished directly in terms of a conflict between prosody and syntax, eliminating the need for any complicating of the syntactic structure prior to PF (however see van Oostendorp et al., this volume, for a slightly nuanced approach).

OT has other properties that make it a parsimonious approach to syntax. Every optimal output is a last resort or the last candidate standing (Grimshaw 2013). Last resort follows from the very architecture of the theory − competition and constraint interaction.

It is never an added principle/constraint but rather a theorem of OT. In her 2013 typology of *do*-support Grimshaw, for example, discusses the fact that a given language may avail itself of alternative last resort strategies in different syntactic contexts. The Northern Italian dialect Monnese (Benincà and Poletto 2004) in (16) is contrasted with German in (17), and with English where *do* support is required in both negative and interrogative contexts (in the absence of an auxiliary).

(16)   a. I so   mìa /*Fo mìa save-l

            It I=know not / I=do not know-it

            'I do not know it'

         b. Ke   fa-l   majà? /*Ke majà-l?

            what does-he eat? / what eats-he?

            'What does he eat?'

(17)   a. Ich weiss es nicht /*Ich tue (es) nicht wissen

             I   know it not /   I do (it) not   know

            'I do not know it'

         b.Was isst er? /*Was tut er essen?

            What eats he? / what does he eat?

            'What does he eat?

         c. Tanzen tut Katja immer noch häufig

             dance does Katja still   often

Monnese opts for V-raising rather than an expletive auxiliary strategy when movement to T is at issue (16a), but opts for the expletive strategy when movement to C is at stake (16b). In contrast, German opts for V-raising in both movement to T and C (17a,b), but an expletive strategy with VP-preposing (17c). The surface form of the expletive element (e.g. *do*) itself is language-particular but its surfacing in a given syntactic context follows from the interaction of universal constraints and their language-particular ranking. Likewise, economy of structure is a theorem of the theory of phrase structure (Grimshaw 2001) and economy of movement follows from the theory of what a chain is, replacing the earlier STAY constraint (Grimshaw 2006). Grimshaw's recent investigations illustrate the evolution of the OT program of research, which seeks to scrutinize earlier OT proposals and achieve increased explanatory depth by eliminating constraints and deriving their effects as consequences of more general principles of the theory.

Parameters circumvent inviolability in MP and, as a consequence, turning off a principle/constraint in a given language *should* result in the complete absence of the principle's active effects. This is not always what is observed empirically. To take one example, German requires expletive *es* in impersonal passives under broad focus or when the adverb denotes new information (Legendre 2001: 9). When new information is also noteworthy/unexpected, *es* does not surface (small caps, where present, represent stress on the focused element).

(18)   Was geschah?

         'What happened?

         Es wurde schön getanzt

'It was beautifully danced'

(19)   Wie wurde getanzt?

         'How was the dancing?

         a. *Es wurde SCHÖN getanzt*

            'It was beautifully danced'

         b. *SCHÖN wurde getanzt.*

            'Beautifully was danced'

German sentences typically contain a subject but the fact that German has an expletive

subject pronoun (*es*) does not entail that it is always optimal. Introducing new and

noteworthy information in German results in a grammatical impersonal passive without a

subject (19b). Under different information structure conditions, an impersonal passive

sentence surfaces with a subject (18, 19a). In the latter case the relatively low-ranked EPP

constraint eliminates a subject-less competitor. The general phenomenon illustrated by

(18)-(19) − a direct consequence of the resolution of conflicting constraints across inputs

which may for example vary by virtue of an information structure property− is known as

The Emergence of the Unmarked (TETU, see McCarthy 2002: 129-34 for further

discussion). The unmarked pattern in German (a sentence with a subject) emerges from

the decisive effect of a low-ranked markedness constraint when dominating constraints

are controlled for (e.g. when an information feature is absent from the input, the relevant

constraint is vacuously satisfied). TETU follows from the very architecture of OT (the

universality of constraints, their rankings, and the EVAL procedure), and TETU effects

are pervasive in OT analyses. In other words, OT *demotes* constraints rather than turning them off. TETU highlights the status of OT as a theory of markedness. OT formalizes what it means to be a marked structure − a structure that is grammatical, but violates (lower-ranked) markedness constraints.

In sum, OT is a *general* theory of grammar, one based on a particular theory of constraint interaction and a theory of the economy of constraint violation. The issue of the representations on which constraints operate (e.g. building blocks, levels of representation, representational vs. derivational architecture) is *orthogonal* to that of constraint interaction. Therefore, OT defines a class of grammatical theories that are as diverse as other theories unified under other umbrellas, e.g. Principles & Parameters theories. These OT theories eschew the need to stipulate special principles of economy, relativized locality, last resort, and the like because these effects follow as inevitable logical consequences of the general competitive architecture.

1.2.3 OT semantics and pragmatics

The optimization-based view on grammar that was originally developed in the domains of phonology and syntax in the early 1990's quickly spread to semantics and pragmatics, in particular in the Dutch research community (Hendriks and de Hoop 1997, 2001; de Hoop and H. de Swart 2000; Blutner and Zeevat 2004; Beaver 2004; Blutner, de Hoop and Hendriks 2006). The realization that competition in form and meaning often go hand in hand led to the development of *bidirectional* OT (Blutner, 2000, 2004; Beaver and Lee 2004; Jäger 2004; Mattausch 2007; Hendriks et al. 2010), as will be outlined in Section 1.3. For now, we concentrate on unidirectional approaches.

OT semantics encodes the interpretation process as the hearer's task of selecting

the optimal interpretation for a particular form used by the speaker. According to

Hendriks and de Hoop (1997, 2001), GEN freely generates candidate meanings. With

Zeevat's F-INT and *INVENT combined, many irrelevant interpretations are weeded out.

But more constraints are needed to account for specific interpretive effects, such as the

complementarity between anaphors and pronouns illustrated in (11) above, repeated here

as (20):

(20) a. John$_i$ saw him$_{*i/j}$/himself$_{i/*j}$ in the mirror.

      b. John$_i$ thinks everyone adores him$_{i/j}$/*himself.

According to Hendriks and de Hoop (2001), the reflexive meaning of the pronoun in

(20a) is blocked because of the interaction between the two constraints in (21a) and

(21b):

(21) a. *DOAP*: Do not overlook anaphoric possibilities

      b. *Principle B*: If two arguments of the same relation are not marked as being

          identical, interpret them as being different.

      c. *World Knowledge*: interpretation fits world knowledge

The constraint ranking *Principle B >> DOAP* blocks the reflexive interpretation in (20a),

as desired. Given that the pronoun and its antecedent are not arguments of the same

relation, *Principle B* is vacuously satisfied in (20b), and an anaphoric interpretation is

preferred. This does not imply that the disjoint interpretation is not available in (20b), because anaphora resolution is a more complex process that brings in other constraints, like the pragmatic constraint *World Knowledge* in (21c). The relevance of this constraint has already been illustrated in (9) and (10), repeated here as (22):

(22) a. Mitt narrowly defeated Rick, and the press promptly followed him to the next

primary state.   [him = Mitt]

　　b. Mitt narrowly defeated Rick, and he quickly demanded a recount.   [he = Rick]

De Hoop (2013) proposes an OT analysis in which Kehler and Rohde's (2013) probabilistic account of the contrast in (22) is modeled in terms of conflicting constraints, with the constraint requiring the interpretation to fit world knowledge outranking a constraint on satisfaction of implicit causality, itself outranking a constraint on topic continuation. As a result, optimization leads to a single, optimal interpretation of the pronoun in context.

For OT approaches to be successful in semantic theory, it is important to investigate the relation between OT and the principle of compositionality of meaning. The principle of compositionality of meaning holds that the meaning of a complex whole is a function of the meaning of its composing parts and of the way these parts are put together by syntactic rules. This principle has immediate appeal, especially if we combine it with function application as the main mode of composition. In a syntactic structure A with two branches B and C, one branch (say B) is interpreted as the functor , and the other one (C) as its argument, so $[[A]] = [[B]]([[C]])$. OT is often taken to be incompatible

with compositionality of meaning, because function application is seemingly incompatible with the notion of competition and variation that drives approaches in terms of optimization. However, optimization complements compositionality when function application does not provide enough information to determine interpretation. The interpretation of adjectives provides an illustration.

The difference between *red apple* (=red peel) and *red grapefruit* (=red pulp) constitutes a well-known example of the problem of adjectival modification (Quine 1960). The problem can be approached from two different theoretical perspectives. One assumes semantic underspecification augmented with contextual enrichment (Bierwisch and Schreuder 1992; Blutner 2004). The other assumes the strongest meaning hypothesis, which leads to overspecification of semantic features in lexical entries and allows weakening in context (Zwarts 2004; Hogeweg this volume; Hoek and de Hoop this volume). Both approaches rely on pragmatics, world knowledge, and some form of optimization process over lexical, syntactic, and contextual information to fix the interpretation of a complex whole beyond function application. Other compositionality puzzles require a bidirectional perspective, as illustrated in Section 1.3 below.

As observed in Section 1.1.2 above, optimization effects in semantics/pragmatics are often intertwined with syntax. For instance, markedness of meaning generally correlates either with morpho-syntactic markedness in pairs of nouns referring to male/female referents, or with the relation between adjectives like *happy* and its double negative *not unhappy*. So it is not surprising that the semantics/pragmatics literature was the first to move from unidirectional to bidirectional optimization. As we will see in

Section 1.3, bidirectional OT quickly spread from semantics to language acquisition and typology, and it proves to be relevant in syntax as well.

1.3 From unidirectional to bidirectional OT

Optimization is direction-sensitive in OT: from meaning to form in *production* (OT syntax), and from form to meaning in *comprehension* (OT semantics). Faithfulness in phonology − as well as faithfulness in purely syntactic optimizations of the type discussed in Legendre et al. (1998) − is defined in terms of maximal identity between input and output. Obviously, forms and meanings are not identical, so faithfulness constraints in OT syntax/semantics can alternatively be seen as driving maximal *correspondence* between forms and meanings. That is, if a particular property of meaning is present in the input, then a particular property of form must be present in the output, and vice versa. Bidirectional approaches have also been developed in phonology (see Boersma 2011), but the fact that forms and meanings cannot be identical, makes the directionality in OT syntax/semantics even more important and raises the question whether speakers take the hearer's perspective into account in production, and whether hearers take the speaker's perspective into account in comprehension. The answer to this question requires a change in perspective from *unidirectional* to *bidirectional* optimization.

Section 1.3.1 presents the essence of strong and weak bidirectional OT as governing the syntax-semantics interface. Section 1.3.2 discusses some implications of bidirectional OT for compositionality of meaning and semantic variation. Section 1.3.3 presents a brief outline of the move from unidirectional to bidirectional approaches in OT

approaches to language acquisition. Besides the symmetric version presented in Section

1.3.1, there are several asymmetric versions of bidirectionality, which are relevant to

analyses of language change (Section 1.3.4). Section 1.3.5 concludes this overview of

syntactic and semantic developments in Optimality Theory, before moving on to a

presentation of the contents of the volume in Section 1.4.

1.3.1 Strong and weak bidirectional Optimality Theory

Bidirectional OT historically emerged as a combination of radical pragmatics and OT

semantics in Blutner (2000). For Grice (1975), the cooperation principle drives the

system of conversational implicatures. Blutner recasts this in terms of the interaction

between speaker and hearer optimization processes. Blutner (2000, 2004) develops a

strong and a weak version of bidirectional OT (biOT) to model complete and partial

blocking phenomena, respectively. Strong bidirectional optimization is defined in (23):

(23)   Strong bidirectional optimization (adapted from Blutner 2000):

        A form-meaning pair <f, m> is bidirectionally optimal iff:

        a.   there is no other pair <f', m> such that <f', m> is more harmonic than

                <f,m>.

        b.   there is no other pair <f, m'> such that <f, m'> is more harmonic than

                <f,m>.

Bidirectional OT is *simultaneous* optimization of production (choice of optimal form for

a certain meaning in 23a) and comprehension (choice of optimal meaning for a certain

form in 23b). As a simple example of *strong* bidirectional optimization, consider the biOT analysis of (24a) and (24b) in Hendriks (2014). Hendriks is interested in the distribution of meanings of anaphors and pronominals in a setting with two participants:

 (24)   a.   The elephant is hitting himself.

              HIT(elephant, elephant)

        b.   The elephant is hitting him.

              HIT(elephant, alligator)

The production and comprehension of sentences with pronouns and reflexives is governed by the two independently motivated constraints in (25):

(25)   a.   *Principle A*: A reflexive must be bound in its local domain.

        b.   *Referential economy*: *lexical NP >> *pronoun >> *reflexive.

Principle A in (25a) is construed as a faithfulness constraint in Hendriks (this volume). It maps forms onto meanings in a way familiar from the original Binding Theory (one important difference is the absence of a Principle B, as discussed in Section 1.2.3. The referential economy scale in (25b) is a markedness constraint hierarchy on forms adapted from Burzio (1998). The following strong bidirectional OT tableau derives the intended form-meaning pairs in (24) as optimal:

[INSERT TABLEAU 1 HERE]

Tableau 1 looks a bit different from a unidirectional tableau, because it requires comparison of candidates consisting of form-meaning pairs. *Referential Economy* favors reflexives over pronouns. *Principle A* relates the bound interpretation to the reflexive form. The pair <the elephant is hitting himself, HIT(elephant, elephant)> constitutes an optimal form-meaning pair (conventionally identified as ✌), because the reflexive is a better *form* for this meaning than the pronoun (given *Referential Economy*), and the bound interpretation is a better *meaning* for this form than disjoint reference (given *Principle A*).

Referential Economy disfavors the use of pronouns, but neither this constraint nor *Principle A* restricts the *interpretation* of pronouns. Hence, pronouns would be ambiguous in a *unidirectional* comprehension grammar consisting of these constraints. By contrast, a bidirectional grammar pairs up the pronoun with the disjoint reference interpretation (see the candidate set in Tableau 1) because the bound interpretation is preferred for the reflexive, and *Principle A* outranks *Referential Economy*.

The notion of strong bidirectional optimality is equivalent to a Nash equilibrium in game theory (Dekker and van Rooy 2000). Under the view that language is a game played by a speaker and a hearer, the complementary distribution of forms and meanings in bidirectional Tableau 1 instantiates a stable communicative strategy that speaker and hearer converge on.

However, not all blocking effects can be accounted for in a strong version of bidirectional OT. Section 1.2.2 introduced the *partial* blocking of *cause to die* by *kill* as an instance of Horn's division of pragmatic labor. Both forms exist, and are distinct in meaning, even though this does not follow in a straightforward way from lexical

semantics. Blutner (2000, 2004) defines a *weak* version of bidirectional optimization, which allows candidates that did not pass the first round of optimization to compete again in a second round, leading to super-optimality:

(26) Weak bidirectional optimization (adapted from Blutner 2000):

A form-meaning pair <f, m> is super-optimal iff:

a. there is no super-optimal pair <f', m> such that <f', m> is more harmonic than <f, m>.

b. there is no super-optimal pair <f, m'> such that <f, m'> is more harmonic than <f, m>.

Blutner (2000) takes the shorter form *kill* to be unmarked relative to periphrastic *cause to die*, and takes the meaning of direct causation to be unmarked relative to indirect causation. Based on markedness, strong bidirectional OT yields the unmarked pair <*kill*, direct causation>, as the best form-meaning combination, and blocks all other pairs as suboptimal. Under the weaker definition of bidirectional optimization in (26), <*cause to die*, indirect causation> emerges as a second super-optimal pair, because there is no better way to express this meaning and no better meaning for this form. Weak bidirectional optimization patterns can be captured in diagrams, where arrows represent the preference relations:

[INSERT TABLEAU 2 HERE]

The horizontal arrows indicate a preference for unmarked forms over marked forms, whereas the vertical arrows indicate a preference for unmarked meanings over marked meanings. All arrows point to the optimal form-meaning pair <*kill*, direct causation> and away from <*cause to die,* indirect causation>. Yet, in a second round of optimization, <*cause to die*, indirect causation> also emerges as a superoptimal candidate, because there is no super-optimal pair that blocks it. The two candidates <*kill,* indirect causation> and <*cause to die*, direct causation> are not super-optimal because they are both blocked by the other super-optimal pair <*kill*, direct causation>. As a result, they are not in competition with the pair <*cause to die*, indirect causation> anymore.

The interaction of lexical and syntactic negation is a slightly more complex example of Horn's division of pragmatic labor. Blutner (2004) assumes that gradable adjectives denote extremes of a scale, so *happy* denotes ☺, *unhappy* denotes ☹, and there is an intermediate state ☺:

[INSERT FIGURE 1 HERE]

We might expect *not unhappy* to denote the complement of ☹, that is the part covering both ☺ and ☺. But *happy* is a simpler form, so markedness blocks the meaning ☺ for *not unhappy*. Under bidirectional optimization, *not unhappy* is thus restricted to ☺. In this way, two negatives do not make a positive, and are not equivalent to affirmation (cf. Horn 1989/2001 and Section 1.2.3 above).

Bidirectional optimization is making its way into the syntax-semantics interface. For example, Legendre, Smolensky, and Culbertson (this volume) develop a weak

bidirectional optimization of blocking effects in aspectual interpretations of French inchoative verbs, an empirical domain previously analyzed in purely syntactic terms. Their analysis is framed in terms of an interaction between lexical, syntactic, and semantic constraints.

Weak bidirectionality has been challenged from the perspective of online processing. Specifically, recursion runs the risk of infinite regression (Beaver and Lee 2004; Blutner et al. 2006; Hendriks et al. 2010). Solutions include limiting recursion to a finite number of steps (Blutner et al. 2006), or restricting the role of weak bidirectionality to intermediate stages in diachronic/evolutionary modeling (Blutner and Zeevat 2009). Crucially these solutions allow the preservation of the two versions necessary to address the full range of blocking phenomena in natural language.

1.3.2 Implications of bidirectional optimization for compositionality and typology

Bidirectional Optimality plays a role in puzzles of compositionality of meaning and cross-linguistic semantics. Under the assumption that compositional semantics and Gricean principles are stable across languages, it could be expected that languages vary in form, but not in meaning. This would restrict the applicability of principles of optimization in semantics. Yet, the current interest in semantic typologies (Chierchia 1998; Evans 2010; Matthewson 2014) suggests otherwise.

One instance of semantic variation arises when compositional semantics allows more than one interpretation and the choice between these interpretations is driven by the syntax. H. de Swart (2010) analyzes the contrast between double negation and negative concord languages in this way. Expressions like *not, nobody*, and *never* are inherently

negative, but combinations of multiple negative expressions in one sentence may lead to double negation (e.g. Standard English, Standard Dutch) or single negation (e.g. Italian, Greek and other so-called negative concord languages), as illustrated in (29):

(29) a.   *Nessuno* ha   detto *niente*.   [Italian]

           Nobody has said   nothing.

           'Nobody has said anything.'   ¬∃x∃y Say(x,y)

      b.   Nobody has said nothing.   ¬¬∃x∃y Say(x,y)

      c.   Nobody has said anything.   ¬∃x∃y Say(x,y)

The combination of *nessuno* and *niente* in (29a) gives rise to a single negation reading that would be conveyed in English by the combination of *nobody* and *anything* in (29c), rather than the combination of *nobody* and *nothing* in (29b). Polyadic generalized quantifier theory derives the double negation reading though iteration of negative quantifiers, and the single negation reading through resumption (de Swart and Sag 2002).

However, polyadic generalized quantifier theory does not tell us which reading to derive. Building on de Swart and Sag (2002), de Swart (2010) uses the conflict between the markedness constraint *NEG and the faithfulness constraint FNEG (be faithful to negation) to derive two types of languages: the ranking *NEG >> FNEG leads to negative concord, and FNEG >> *NEG to double negation. Thus two possible interpretations are available through compositional semantics, and OT derives the optimal choice through the usual means − constraint ranking. Ambiguities are rare though, so the choice between the two operations is clearly grammaticized. Languages vary in the ranking they assign to

MAXNEG, the constraint that multiplies the realization of negation on all the arguments. Negative concord languages then have the bidirectional ranking MAXNEG >> *NEG >> FNEG, whereas double negation languages have the bidirectional ranking FNEG >> *NEG >> MAXNEG. Compositional interpretation is thus defined bidirectionally.

1.3.3 From unidirectional to bidirectional OT in studies on language acquisition

Early OT applications to the problem of language acquisition have focused on child production, and are thus typically framed in a unidirectional grammar. *Optionality* in child production, whereby the forms young children produce typically alternate between adult-like and non-adult-like sentences for a while, also illustrates the challenge of accounting for (true) optionality in an optimization-based system.

Consider the stage of linguistic development dubbed the Optional Root Infinitive stage in Wexler (1994) during which young children around the age of 2 produce both tensed/agreeing sentences (*He is going there*) as well as non-tensed/agreeing ones (**He go there*) in a number of languages. Wexler (1998), for example, proposes a model, which relies on adding child-specific constraints to child grammars, which eventually disappear though the (unformalized) mechanism of maturation. An alternative OT model of child production (e.g., Legendre et al. 2002, 2004; Davidson and Legendre 2003) appeals instead to the partial ranking of "floating" constraints originally proposed to model dialectal variation (Anttila 1997; Nagy and Reynolds 1994) and straightforwardly accounts for both the optionality of syntactic phenomena in child language (e.g. root infinitives in French vs. Catalan) and the frequency with which alternative forms evolve over time until an adult-like grammar of the relevant syntactic phenomena is reached.

Extending an optimization-based approach to multilingual grammars represents an empirical domain that to date has been under-explored. In light of recent psycholinguistic research that suggests that bilinguals have the ability to simultaneously activate multiple grammars (e.g. Kootstra et al. 2010; Kroll et al. 2006; Kroll et al. 2012), some systematic predictive formalism must be in place to model the interaction of these two grammars in the bilingual mind. To this end, Bousquette et al. (this volume) sketch how a model employing floating constraints and partial constraint rankings (as well as Engdahl's 1983/2001 accessibility hierarchy) converge on accounting for the parasitic gap pattern found in a heritage variety of German spoken in Eastern Wisconsin, which is distinct from parasitic gap patterns in both English and Standard German. Another example is Hsin (2014) who formalizes an integrated early bilingual grammar architecture in OT and accounts for both facilitation and interference effects across two languages (English/Spanish) acquired simultaneously. Both proposals involve unidirectional OT analyses that pertain to production.

But children also need to acquire the interpretation principles of their mother tongue, so in this domain as well, a move from a unidirectional perspective to a bidirectional one is in place. For example, an early sketch of full acquisition, e.g. simultaneous adult-like comprehension and non-adult-like production, in terms of dual optimization is made in Legendre et al. (2004) in the context of the acquisition of Mandarin aspectual particles. But comprehension does not necessarily precede production in acquisition and detailed studies of the reverse pattern have led to the general claim that moving from unidirectional to bidirectional optimization constitutes a

crucial step in the child's development towards a full mastery of the adult grammar (e.g. Legendre and Smolensky 2012; Hendriks 2014, this volume).

Hendriks' (2014) strong bidirectional OT analysis of the distribution and meanings of anaphors and pronominals discussed above is specifically motivated by her interest in explaining the fact (see e.g. Chien and Wexler 1990) that young children are known to misinterpret the pronoun in (24b), repeated here as (28b), as coreferential with the subject, even though they correctly produce the reflexive in contexts like (24a), repeated here as (28a).

(28)　　a.　　　The elephant is hitting himself.

　　　　　　　　HIT(elephant, elephant)

　　　　b.　　　The elephant is hitting him.

　　　　　　　　HIT(elephant, alligator)

Hendriks (this volume) shows that the asymmetry between production and comprehension is also found with other phenomena. In her explanation, children are unable to optimize bidirectionally, so every time a markedness constraint is ranked too high in the child's grammar relative to a conflicting faithfulness constraint, there will be a mismatch between production and comprehension. If the markedness constraint applies to surface forms, production is delayed compared to comprehension. If the markedness constraint applies to meanings, comprehension is delayed compared to production.

In the case of anaphors and pronominals, Referential Economy disfavors the use of pronouns, but neither this constraint nor Principle A restricts the *interpretation* of

pronouns. Hence, pronouns are ambiguous in a *unidirectional* comprehension grammar consisting of these constraints, which Hendriks argues is characteristic of the child's grammar. In the bidirectional adult grammar, the pronoun is paired up with the disjoint reference interpretation (see the candidate set in Tableau 1), because the reflexive is assigned the bound interpretation. According to Hendriks and Spenader (2006) and Hendriks (2014), children are unable to reason bidirectionally until they have acquired a Theory of Mind.

Both strong and weak bidirectional optimization are *symmetric*, in that they rely on the intuition that speakers take into account the hearer's perspective, and hearers the speaker's perspective. However, some researchers have defended the need for asymmetric versions of bidirectionality, as discussed in Section 1.3.4.

1.3.4 Asymmetric bidirectionality

Zeevat (2001, 2006) develops an asymmetrical version of OT in which a unidirectional OT model for production forms the basic system. In comprehension, the set of candidate meanings is restricted to the outcome of the production-directed optimization. The opposite view has been defended by Jäger (2004), whose asymmetric model discards forms as optimal candidates when they are not optimally recoverable as the intended meaning and at least one other candidate is. Formally:

(27)   i.   A form-meaning pair <f,m> is hearer-optimal iff there is no pair <f,m'> such that <f,m'> is better than <f,m>;

        ii.   A form-meaning pair <f,m> is optimal iff <f,m> is hearer-optimal, and

there is no distinct pair <f',m> such that <f',m> is better than <f,m> and

<f',m> is hearer-optimal, or no pair is hearer-optimal, and there is no

distinct pair <f',m> such that <f',m> is better than <f,m>.

In Jäger's system, speakers take hearers into account when they formulate their message, but hearers do not take into account speakers when they interpret. In combination with a learning algorithm, Jäger shows how this asymmetric version of biOT accounts for language change and evolution.

The diachronic analysis of differential object marking developed by Lestrade, van Bergen and P. de Swart (this volume) relies on a Jäger style asymmetric bidirectional OT. They ground this in two general principles: the principle of Economy (least effort), and the drive towards communicative success (the speaker wants to be understood by the hearer). Lestrade et al. include a principled discussion of the relation between unidirectional and bidirectional OT and show how the standard unidirectional (syntactic) analysis of differential object marking in the literature can be understood as the combination of asymmetric bidirectional optimization and a mechanism that generalizes over frequent patterns. Overall, their paper illustrates the need for bidirectional mechanisms in analyzing patterns of diachronic change and grammaticalization.

1.3.5 Conclusion

We conclude our presentation of OT's specific properties and most recent developments by reflecting on the goals for linguistic theorizing that were articulated by Lees (1957: 376) in the very early days of generative grammar, namely:

(30) a. Freedom from contradiction,

  b. Maximal cohesion with other branches of knowledge,

  c. Maximal validity in coverage of known data, and

  d. Maximal elegance of statement.

In our view, (30) remains a model for capturing the relationship between raw, natural data and any sort of axiomatic symbolic logic used to model its inherent core principles. In the preceding sections, we have sought to make the case that OT instantiates Lees' desiderata by highlighting three points: (1) OT permits an architectural unification of syntax with semantics/pragmatics (not to mention phonology) thanks in particular to developments in bidirectional OT, and provides novel insights on how these two core modules interact; (2) OT can profitably be construed as a theory of the interfaces with other modules, e.g. prosody; and (3) a grammar architecture grounded in comparison and evaluation simplifies the overall theory of the linguistic faculty.

  While modeling the grammar as an optimizing system, specifically as formulated in OT, can be seen as an attempt to emulate Lees' high standards for linguistic theorizing our position should not be misunderstood as one that finds no challenges for an OT approach to modeling human grammars. For example, OT has had to confront structural ineffability (see Section 1.2 and the discussion of input-output faithfulness), opacity, compositionality in semantics (see Section 1.3), etc.[i] If anything, the sequence of alternative models proposed by Chomsky himself since the 1950's have taught us that

trade-offs must typically be made, among theoretical desiderata and/or empirical coverage.

1.4. Content of the volume

The chapters in this volume are, for the most part, based on talks presented at a workshop organized at Johns Hopkins University by Géraldine Legendre and Mike Putnam in November 2012. They focus on both fundamental and targeted issues in syntax (Part I) and in semantics/pragmatics (Part II). Collectively, they illustrate the independence of a theory of constraint interaction from the substantive theory, which provides particular statement of these constraints. The first six papers investigate a variety of syntactic phenomena and in some cases their interface with other components of the grammar. The remaining six papers investigate a variety of semantic/pragmatic phenomena and highlight the need for bidirectionality of optimization.

Part 1: Syntax and interfaces

Marc van Oostendorp, Michael Putnam, and Laura Catharine Smith engage in a discussion of families of violable constraints in their contribution entitled, "Intersecting constraints - Why certain constraint-types overlap while others don't." From an architectural standpoint, van Oostendorp et al. make the case for a parallel model of grammar over modular alternatives. To illustrate their point with respect to the syntax-phonology interface, van Oostendorp et al. not only make the case for a parallel architecture of grammar, but also call for a reassessment of constraint well-formedness in this area of linguistic investigation. Van Oostendorp et al. introduce a restrictive theory of

what a plausible OT constraint could look like at the boundary of syntax and phonology, suggesting that "structural constraints" in syntax (i.e. phrase structure rules) and in phonology (i.e. syllable structure) can, in fact, overlap, thus reducing the inventory of constraints while exponentially increasing their power. This view of constraints, which focuses on the interaction of multiple levels of grammar, provides an alternative theory of "interfaces" to those found in mainstream generative grammar (cf. Reinhart 2006; Broekhuis and Vogel 2013).

In his contribution entitled "Feature inheritance versus extended projections", Hans Broekhuis revisits recent theoretical claims in the Minimalist literature that assimilates phi-features values on complementizers (C) and heads of tense phrases (T) under one single operation known as FEATURE INHERITANCE (see e.g. Chomsky 2008; Pesetsky and Torrego 2001, 2004, 2007). Broekhuis challenges recent analyses of subject vs. object *wh*-extractions that adopt the notion of phases and Feature Inheritance, and in their place, provides argumentation that Grimshaw's (1997) formulation of extended projections in addition to economy/optimization principles results in a simpler, more elegant theoretical analysis. The paper offers support for the core concept of minimal violation in syntax. Specifically, CP-structures can be harmonically bounded by (and lose to) TP-structures due to the fact that they invoke additional violations of derivational operations such as *MOVE/INTERNAL MERGE and *MERGE/EXTERNAL MERGE.

Fabian Heck and Gereon Müller also address the ordering of the Minimalist derivational operations MERGE and MOVE in their contribution, "On Accelerating and decelerating movement: From Minimalist preference principles to harmonic serialism." Building on Assmann et al. (in press), which focuses on the competition between Agree

and Move, Heck and Müller provide evidence that the derivational structure building principle of "Merge before Move" (cf. Chomsky 2000) is untenable. They argue for subsuming Probe-Goal relations (= Agree) under Merge, and show that Agree must procrastinate in favor of Move or vice versa, depending on whether a language has an accusative or ergative case system. In turn, procrastination of Agree suggests constraint violability, which Heck and Müller argue must exist to arrive at a typologically-valid theoretical account of basic structure-building operations. In this regard, Heck and Müller's paper returns us to original arguments laid out by Speas (1997) questioning the inviolability of structure-building operations in a derivational syntactic framework.

In "Two Types of Portmanteau Agreement: Syntactic and Morphological" Ellen Woolford argues against lumping together all instances of portmanteau agreement cross-linguistically. She proposes to distinguish *syntactic* portmanteau agreement, an instance of multiple Agree relations in the syntax that is restricted to ergative languages, from *morphological* portmanteau agreement in accusative languages, which involves optimization over violable alignment and faithfulness constraints at PF. This contribution serves as an additional example of the mutual benefit of invoking constructs from both the Minimalist Program and from OT.

In "Optimal constructions", Ralf Vogel explores the potential benefits of integrating violable, ranked constraints into a version of Construction Grammar (CxG; Goldberg 2006). More specifically, Vogel explores how optimization achieved through violable OT-constraints is necessary to determine variable argument linking in CxG. Vogel argues that due to the underspecified nature of constructions that require some sort of linking algorithm, neither the lexical entry of the verb nor the larger unit labeled as a

'construction' are typically sufficient to determine successful argument linking. He sketches out a theoretical system where contextual knowledge as well as these aforementioned units can be mediated through violable constraints ranked in an OT-grammar. This contribution illustrates that optimization is independent of the particular substantive theory of grammar (its representations, operations, etc.) it is applied to.

Joshua Bousquette, Michael Putnam, Joseph Salmons, Benjamin Frey, and Daniel Nützel investigate the potential conceptual advantages of modeling heritage grammars, first-language attrition, and language change from an optimality-theoretic perspective in their contribution, "Multiple grammars, dominance and optimization." In their study, Bousquette et al. examine production data from 41 speakers of Wisconsin Heritage German (WHG), with the empirical focus being on their ability to produce multiple filler-gap dependencies (e.g. parasitic gaps), which are structures that are highly marked in standard and non-standard German. Bousquette et al. employ floating constraints (e.g. Anttila 1997; Legendre et al. 2002) making use of a power constraint hierarchy based on Engdahl's (1983/2001) Accessibility Hierarchy to account for their empirical findings, suggesting that first-language attrition has played a relatively minimal role in the licensing of parasitic gaps in WHG. Bousquette et al.'s contribution opens the door to ongoing productive research in the domains of bi- and multilingualism in the area of first-language attrition, grammar systems across the lifespan, and its connection with research conducted from an OT-perspective in monolingual and bilingual acquisition (e.g. Hsin 2014).

Part 2: Semantics/Pragmatics

In "On the origin of constraints", Sander Lestrade, Geertje van Bergen, and Peter de Swart address two foundational issues in OT: whether constraints are innate or not and whether optimization is unidirectional or bidirectional. As a case study, they investigate the development of constraints on object marking and the development of differential case marking systems. They show through computer simulation how, by generalizing over the outcomes of bidirectional optimization procedures in which the hearer's perspective is considered, a speaker learns that certain meanings require additional marking for successful communication. Then, once the correlation between input meaning and optimization outcome reaches a critical threshold, a direct link between meaning and form will be established. The authors thus show that unidirectional constraints on object marking need not be innate and can be the result of the fossilization of bidirectional optimization processes.

In "Optimality Theory and lexical interpretation and selection" Lotte Hogeweg discusses cases where word meanings are flexible (e.g. *fast* in *fast car* vs. *fast food*) but where there are no straightforward rules to determine their meaning in context. She takes up the issue of which aspects of meaning should be specified in the lexicon and how these should be altered in a given context. She argues for an 'overspecified' approach, where all of a word's semantic features become activated when that word is encountered, and features can subsequently be deactivated through a process of optimization. She further demonstrates that this overspecification approach is supported by psycholinguistics data.

Jet Hoek and Helen de Hoop highlight the surprising difficulty in determining the appropriate interpretation for Dutch *ja* 'yes' and *nee* 'no' in their contribution entitled

"On the Optimal Interpretation of *Yes* and *No* in Dutch". For example, a negative

response *nee* can indicate surprise and unhappiness without indicating disagreement, and

it can even be used to agree with a negative statement. By examining a corpus of spoken

Dutch, the authors identify a number of distinct possible interpretations for *ja* and *nee*,

which they arrange along a hierarchy of strength, measured in terms of referential

distance and concreteness. While stronger meanings are generally preferred, they show

that this preference can be overruled in the interest of choosing a meaning that fits the

context. They go on to provide a bidirectional OT account of both the Dutch and English

answering systems, where appropriate utterance-response pairs are determined by the

relative ranking of constraints on aligning the polarity of utterances, avoiding negation,

mapping *ja*/*yes* to positive interpretations and *nee*/*no* to negative interpretations, and

economy.

In her contribution entitled "Telicity features of bare nominals", Henriette de

Swart seeks to explain the cross-linguistic distribution and interpretation of bare

nominals. This issue is particularly challenging given that bare nominals do not lend

themselves to a compositional analysis the way nominals appearing with determiners do.

She questions the extent to which the distribution of bare nominals is part of their

universal semantics or due to cross-linguistic considerations. By examining a number of

languages, she argues that their distribution within- and cross-linguistically is determined

by the relative ranking of constraints which also determine the language's number

morphology and article system. She relates features such as telicity, atomic reference, and

definiteness under a blocking analysis in bidirectional OT with a subset of language-

specific rankings to account for the interpretations of various nominals across a number of different languages.

In "Blocking effects at the lexicon/semantics interface and bi-directional optimization in French" Géraldine Legendre, Paul Smolensky, and Jennifer Culbertson investigate French inchoative verbs that divide into three subclasses based on whether the reflexive marker *se* is prohibited, required, or optional. They demonstrate that variability in the interpretation of subject pronoun *il* (referential vs. non-referential), as well as in aspectual interpretation (result-oriented or not) does not line up with morphologically defined subclasses. The authors show that, while problematic under standard (morpho)-syntactic accounts, this variability can be straightforwardly accounted for when viewed as resulting from blocking effects between form-meaning pairs in a bidirectional OT framework. For example, a non-referential interpretation is available, except where a globally preferred referential one is available (as is the case for the ambiguously transitive/intransitive verb *casser* 'break'). In this analysis, lexical gaps, markedness constraints, and competition among expression/interpretation pairs conspire to produce a mismatch between inchoative subclasses and their interpretation in particular sentential contexts.

Petra Hendriks investigates asymmetries that arise between children's language production and comprehension in her contribution entitled "Unfaithful conduct: A competence-based explanation of asymmetries between production and comprehension". This can be seen, for example, in children's mispronunciation of words that they correctly understand (e.g. a child may correctly understand /kæt/ as 'cat', but may produce it as /ta/), as well as in their misunderstanding of utterances they correctly

produce (e.g. SVO production but OVS comprehension of sentences like "The car is pushing the boy"). She argues for a competence-based account provided by OT, a framework where mappings from form to meaning need not match mappings from meaning to form. She identifies two types of constraint configurations that allow these asymmetries: faithfulness constraints that are dominated by competing constraints, and faithfulness constraints underspecifying a form-meaning mapping.

References

Anderson, John R. (1990). *The Adaptive Character of Thought*. Hillsdale, NJ: Erlbaum.

Anttila, Arto (1997). *Variation in Finnish Phonology and Morphology* (Unpublished doctoral dissertation). Stanford University, Stanford, CA.

Assman, Anke, Doreen Georgi, Fabian Heck, Gereon Müller, & Philipp Weisser (in press). 'Ergatives move too early', *Syntax*.

Beaver, David (2004). 'The optimization of discourse anaphora', *Linguistics and Philosophy,* 27: 3-56.

Beaver, David and Hanjung Lee (2004). 'Input-output mismatches in OT', in R. Blutner and H. Zeevat (eds.), *Optimality Theory and Pragmatics*. Basingstoke, England: Palgrave Macmillan, 112-53.

Benincà, Paola, and Cecilia Poletto (2004). 'A case of do-support in Romance', *Natural Language & Linguistic Theory,* 22(1): 51-94.

Bierwisch, Manfred & Robert Schreuder (1992). 'From concepts to lexical items', *Cognition,* 42: 63-60.

Blutner, Reinhard (2000). 'Some aspects of optimality in natural language interpretation', *Journal of Semantics,* 17, 189-216.

Blutner, Reinhard (2004). 'Pragmatics and the lexicon', in L. Horn and G. Ward (eds.), *Handbook of Pragmatics*. Oxford: Blackwell, 488-514.

Blutner, Reinhard, Helen de Hoop, and Petra Hendriks (2006). *Optimal Communication*. Stanford, CA: CSLI Publications.

Blutner, Reinhard and Henk Zeevat (2004). *Optimality Theory and Pragmatics*. Basingstoke, England: Palgrave Macmillan.

Blutner, Reinhard and Henk Zeevat (2009). 'Optimality-theoretic pragmatics', *ZAS working papers in Linguistics,* 51: 1-25.

Bobaljik, Jonathan (1995). *Morphosyntax: The Syntax of Verbal Inflection* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Boeckx, Cedric (2014). *Elementary Syntactic Structures: Prospects of a Feature-Free Syntax*. Cambridge: Cambridge University Press.

Boersma, Paul (2011). 'A programme for bidirectional phonology and phonetics and their acquisition and evolution', in D. E. Holt (ed.), *Optimality Theory and language change*. Dordrecht: Kluwer, 31-65.

Borer, Hagit (2013). *Taking Form - Structuring Sense Vol. III*. Oxford: Oxford University Press.

Bousquette, Joshua, Michael Putnam, Joseph Salmons, Ben Freyd & Daniel Nützele (this volume). 'Multiple grammars, dominance, and optimization'.

Broekhuis, Hans & Ralf Vogel (2013). *Linguistic derivations and filtering: Minimalism and Optimality Theory*. London: Equinox.

Büring, Daniel (2007). 'Intonation, semantics and information structure', in G. Ramchand & C. Reiss (eds.), *The Oxford Handbook of Linguistic Interfaces.* Oxford: Oxford University Press, 445-74.

Burzio, Luigi (1989). 'On the non-existence of disjoint reference principles', *Rivista di Grammatica Generativa*, 14: 3-27.

Burzio, Luigi (1991). 'The morphological basis of anaphora', *Journal of Linguistics,* 27(1): 1-60.

Burzio, Luigi (1998). 'Anaphora and soft constraints', in P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky (eds.), *Is the Best Good Enough? Optimality and Competition in Syntax*. Cambridge: MIT Press, 93-113.

Carlson, Gregory N. (1980). *Reference to Kinds in English.* New York: Garland Publishing.

Chien, Yu-Chin and Kenneth Wexler (1990). 'Children's knowledge of locality conditions on binding as evidence for the modularity of syntax and pragmatics', *Language Acquisition,* 1: 225–95.

Chierchia, Gennaro (1998). 'Reference to kinds across languages', *Natural Language Semantics,* 6: 339-405.

Chomsky, Noam (1973). 'Conditions on Transformations', in S. R. Anderson and P Kiparsky (eds.), *A Festschrift for Morris Halle.* New York: Holt, Rinehard & Winston, 232-86.

Chomsky, Noam (1977). 'On wh-movement', in P. W. Culicover et al. (eds.), *Formal Syntax*. New York: Academic Press, 71-132.

Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, Noam (1986). *Knowledge of language: its nature, origin, and use.* New York: Praeger.

Chomsky, Noam (1991). 'Some notes on economy of derivation and representation', in R. Freidin (ed.), *Principles and Parameters in Comparative Grammar*. Cambridge, MA: MIT Press, 41-58.

Chomsky, Noam (1993). 'A minimalist program for linguistic theory', in K. Hale and S. J. Keyser (eds.), *The View from Building 20*. Cambridge, MA: MIT Press, 1-52.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam (1998). *Minimalist inquiries: The framework*. MIT Occasional Papers in Linguistics 15. Cambridge, MA: MIT Working Papers in Linguistics.

Chomsky, Noam (2000). 'Minimalist inquiries: the framework', in R. Martin, D. Michaels, and J. Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. Cambridge, MA.: MIT Press, 89-155.

Chomsky, Noam (2008). 'On phases', in R. Freidin, C. P. Otero & M.L. Zubizarreta (eds.), *Foundational issues in linguistic theory. Essays in Honor of Jean-Roger Vergnaud*. Cambridge, MA: MIT Press, 133-66.

Dahl, Östen (1979). 'Typology of sentence negation', *Linguistics,* 17: 79-106.

Dalrymple, Mary, & Kaplan, Ronald M. (2000). 'Feature indeterminacy and feature resolution', *Language*: 759-798.

Davidson, Lisa, & Géraldine Legendre (2003). 'Defaults and competition in the acquisition of functional categories in Catalan and French', in R. Nuñez-Cedeño, L. López & R.Cameron (eds.), *A Romance Perspective on Language Knowledge and Use*. Philadelphia: John Benjamins, 273-90.

Dekker, Paul and Robert van Rooy (2000). 'Bi-directional optimality theory: an application of game theory', *Journal of Semantics,* 17: 217-242.

Diesing, Molly (1996). 'Semantic variables and object shift', *Studies in Comparative Germanic Syntax*, 2: 66-84.

Engdahl, Elisabet (2001). 'Parasitic gaps', in P. Culicover & P. Postal (eds.), *Parasitic gaps*. Cambridge, MA: MIT Press, 69-98. (Reprinted from *Linguistics and Philosophy*, 6: 5-34, 1983)

Evans, Nicholas (2010). 'Semantic Typology', in Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*. Oxford/New York: Oxford University Press, 504-33.

Geman, Stuart, and Donald Geman (1984). 'Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images', *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 6: 721-741.

Givón, Talmy (1979). *On Understanding Grammar*. New York: Academic Press.

Goldberg, Adele (2006). *Constructions at Work*. Oxford: Oxford University Press.

Grice, Paul (1975). 'Logic and conversation', in P. Cole and J. Morgan (eds.), *Syntax and semantics, Vol. 3: Speech acts*. New York: Academic Press, 41–58.

Grimshaw, Jane (1997). Projection, heads and optimality, *Linguistic Inquiry,* 28: 373-422.

Grimshaw, Jane (2001). 'Economy of structure in OT', *Rutgers Optimality Archive*, ROA-434.

Grimshaw, Jane (2006). 'Chains as unfaithful optima', in E. Baković, J. Ito, and J. McCarthy (eds.), *Wondering at the Natural Fecundity of Things: Essays in Honor of Alan Prince*. Linguistics Research Center, University of California, Santa Cruz, 97-109.

Grimshaw, Jane. (2013). 'Last Resorts: A Typology of do-support', in H. Broekhuis and R. Vogel (eds.), *Linguistic derivations and filtering: Minimalism and Optimality Theory*. London: Equinox, 267-95.

Grosz, Barbara, Aravind Joshi & Scott Weinstein. (1995). 'Centering: a framework for modeling the local coherence of discourse', *Computational Linguistics,* 21: 203-225.

Hale, John, and Géraldine Legendre (2004). 'Minimal Links, Remnant Movement, and (Non)-derivational Grammar', in A. Stepanov, G. Fanselow and R. Vogel (eds.), *Minimality Effects in Syntax.* Berlin: Mouton de Gruyter, 177-203.

Haspelmath, Martin (2006). 'Against markedness (and what to replace it with)', *Journal of Linguistics,* 42: 25-70.

Hawkins, John A. (1991). 'On (in)definite articles: implicatures and (un)grammaticality prediction', *Journal of Linguistics,* 27: 405-42.

Heim, Irene (1991). 'Artikel und Definitheit', in A. v. Stechow and D. Wunderlich (eds.), *Handbuch der Semantik.* Berlin: de Gruyter, 487-535.

Hellan, Lars (1988). *Anaphora in Norwegian and the Theory of Grammar.* Dordrecht: Foris.

Hendriks, Petra (2014). *Asymmetries between Language Production and Comprehension.* Dordrecht: Springer.

Hendriks, Petra (this volume). 'Unfaithful conduct: A competence-based explanation of asymmetries between production and comprehension'.

Hendriks, Petra and Helen de Hoop (1997). 'On the interpretation of semantic relations in the absence of syntactic structure', in P. Dekker, M. Stokhof and Y. Venema (eds.), *Proceedings of the 11th Amsterdam Colloquium.* Amsterdam: ILLC, University of Amsterdam, 157-62.

Hendriks, Petra and Helen de Hoop (2001). 'Optimality-theoretic semantics', *Linguistics and Philosophy,* 24: 1-32.

Hendriks, Petra and Jennifer Spenader (2006). 'When production precedes comprehension: An optimization approach to the acquisition of pronouns', *Language Acquisition,* 13(4): 319-48.

Hendriks, Petra, Helen de Hoop, Irene Krämer, Henriette de Swart, and Joost Zwarts (2010). *Conflicts in Interpretation*. London: Equinox.

Hobbs, Jerry R., Mark E. Stickel, Douglas E. Appelt, and Paul Martin (1993). 'Interpretation as abduction', *Artificial Intelligence*, 6: 69-142.

Hoek, Jet and Helen de Hoop (this volume). 'On the optimal interpretation of Y*es* and N*o* in Dutch'.

Hogeweg, Lotte. (this volume). Optimality Theory and lexical interpretation and selection.

de Hoop, Helen (2013). 'Incremental optimization of pronoun interpretation', *Theoretical Linguistics,* 39: 87-93.

de Hoop, Helen & Henriette de Swart (2000). 'Temporal adjunct clauses in Optimality Theory', *Italian Journal of Linguistics*, 12(1): 107-27.

Horn, Laurence (1984). 'Towards a new taxonomy of pragmatic inference: Q-based and R-based implicature', in D. Schiffrin (ed.), *Meaning, Form, and Use in Context: Linguistic Applications/GURT '84*. Washington, D.C.: Georgetown University Press, 11-42.

Horn, Laurence (2001). *A Natural History of Negation*. Chicago: University of Chicago Press. (Original work published 1989).

Hsin, Lisa (2014). *Explorations in Integrated Bilingual Grammatical Architecture* (Unpublished doctoral dissertation). Johns Hopkins University. Baltimore, MD.

Jäger, Gerhard (2004). 'Learning constraint sub-hierarchies', in R. Blutner and H. Zeevat (eds.), *Optimality Theory and Pragmatics*. Basingstoke, England: Palgrave MacMillan, 251-287.

Jakobson, Roman (1984). 'Structure of the Russian verb', in L. R. Waugh and M. Halle (Eds.), *Russian and Slavic grammar – Studies 1931–1981*. The Hague: Mouton, 1–14.

Kehler, Andrew and Hannah Rohde (2013). 'Aspects of a theory of pronoun interpretation', *Theoretical Linguistics,* 39: 295–309.

Kiparsky, Paul (1973). '"Elsewhere" in phonology', in S. Anderson and P. Kiparsky (Eds.), *A Festschrift for Morris Halle*. New York: Hold, Rinehart & Winston, 93-106.

Kootstra, Gerrit, Janet van Hell, & Ton Dijkstra (2010). 'Alignment in the production of code-switched sentences: the role of word order and dialogue partner', *Journal of Memory and Language,* 63: 210-36.

Kroll, Judith, Susan Bobb, and Zofia Wodniecka (2006). 'Language selectivity is the exception, not the rule: arguments against a fixed locus of language selection in bilingual speech', *Bilingualism: Language and Cognition,* 9(2): 119-35.

Kroll, Judith, Paola Dussias, Cari Bogulski, & Jorge Valdes Kroff (2012). 'Juggling two languages in one mind: what bilinguals tell us about language processing and its consequences for cognition', in B. Ross (ed.), *The psychology of learning and motivation*. San Diego: Academic Press, 229-262.

Lees, Robert (1957). 'Review of Syntactic Structures', *Language,* 33(3): 375-408.

Legendre, Géraldine (2001). 'Introduction to OT syntax', in G. Legendre, S. Vikner, and J. Grimshaw (eds.), *Optimality-Theoretic Syntax*. Cambridge, MA: MIT Press, 1-27.

Legendre, Géraldine (2009). 'The Neutralization Approach to Ineffability in Syntax', in C. Rice and S. Blaho (eds.), *Modeling Ungrammaticality in Optimality Theory*. London: Equinox, 237-66.

Legendre, Géraldine and Paul Smolensky (2012). 'On the asymmetrical difficulty of acquiring person reference with personal pronouns: From unidirectional to bidirectional OT', *Journal of Logic, Language, and Information,* 21: 7-30.

Legendre, Géraldine, Paul Smolensky, and Jennifer Culbertson (this volume). 'Blocking effects at the lexicon/semantics interface and bi-directional optimization in French'.

Legendre, Géraldine, Paul Hagstrom, Anne Vainikka, and Marina Todorova (2002). 'Partial constraint ranking in child French syntax', *Language Acquisition,* 10(3): 189-227.

Legendre, Géraldine, Paul Hagstrom, Joan Chen-Main, Liang Tao, and Paul Smolensky (2004). 'Deriving output probabilities in Child Mandarin from a dual-optimization grammar', *Lingua,* 114(9-10): 1147-85.

Legendre, Géraldine, Paul Smolensky, and Colin Wilson (1998). 'When is Less More? Faithfulness and Minimal Links in Wh-Chains', in P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky (eds.), *Is the Best Good Enough? Optimality and Competition in Syntax*. Cambridge, MA: MIT Press, 249-89.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky (1990a). 'Harmonic Grammar —A formal multi-level connectionist theory of linguistic well-formedness: Theoretical foundations', in *Proceedings of the twelfth annual conference of the Cognitive Science Society*. Cambridge, MA: Erlbaum, 388-95.

Legendre, Géraldine, Yoshiro Miyata, and Paul Smolensky (1990b). 'Can connectionism contribute to syntax? Harmonic Grammar, with an application', in M. Ziolkowski, M. Noske, and K. Deaton (eds.), *Proceedings of the 26th regional meeting of the Chicago Linguistics Society*. Chicago: Chicago Linguistic Society, 237–52.

Lestrade, Sander, Geertje van Bergen, and Peter de Swart (this volume). 'On the origin of constraints'.

Mattausch, Jason (2007). 'Optimality, bidirectionality & the evolution of binding phenomena', *Research on Language & Computation,* 5: 103-131.

Matthewson, Lisa (2014). 'On how (not) to uncover cross-linguistic variation', in S. Keine and S. Sloggett (eds.), *Proceedings of the 42nd meeting of the North East Linguistic Society*. Amherst, MA: GLSA, 323-42.

McCarthy, John (2002). *A Thematic Guide to Optimality Theory*. Cambridge: Cambridge University Press.

McCawley, James D. (1978). 'Conversational implicature and the lexicon', in P. Cole (ed.), *Syntax and Semantics 9: Pragmatics*. New York: Academic Press, 245-259.

Müller, Gereon (2009). 'Optimality-theoretic Syntax', Lecture notes for a compact course. Unpublished Manuscript.

Müller, Gereon (2011). *Constraints on Displacement: A Phase-based Approach*. Amsterdam: John Benjamins.

Müller, Gereon (2012). *Optimality-Theoretic Syntax*. Unpublished manuscript.

Müller, Gereon and Wolfgang Sternefeld (2001). 'The rise of competition in syntax: A synopsis', in G. Müller and W. Sternefeld (eds.), *Competition in syntax (Vol. 49)*. Berlin: Walter de Gruyter, 1-67.

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

Newmeyer, Frederick (2005). *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.

Oaksford, Mike and Nick Chater (1998). *Rational Models of Cognition*. Oxford: Oxford University Press.

van Oostendorp, Marc, Michael Putnam & Laura Catharine Smith (this volume). 'Intersecting constraints - Why certain constraint-types overlap while others don't'.

Pater, Joe (in press). 'Universal Grammar with weighted constraints' in J. McCarthy and J. Pater (Eds.), *Harmonic Grammar and Harmonic Serialism.* London: Equinox.

Pesetsky, David (1997). 'Optimality Theory and syntax: movement and pronunciation', in D. Archangeli & D.T. Langendoen (eds.), *Optimality Theory: An Overview*. Oxford: Blackwell-Wiley, 134-170.

Pesetsky, David & Ester Torrego (2001). 'T-to-C movement: cases and consequences', in M. Kenstowicz (d.), *Ken Hale, A Life in Language.* Cambridge, MA: MIT Press, 355-426.

Pesetsky, David & Ester Torrego (2004). 'Tense, case, and the nature of syntactic categories' in J. Guéron & J. Lecarme (eds.), *The syntax of Time*. Cambridge, MA: MIT Press, 495-537.

Pesetsky, David & Ester Torrego (2007). 'The syntax of valuation and the interpretability of features' in S. Karimi, V. Samiian and W. Wilkins (eds.), *Phrasal and Clausal Architecture*. Amsterdam: John Benjamins, 262-94.

Phillips, Colin and Shevaun Lewis (2013). 'Derivational order in syntax: Evidence and architectural consequences', *Studies in Linguistics,* 6: 11-47.

Prince, Alan and Paul Smolensky (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell. (Original work published 1993)

Quine, Willard Van Orman (1960). *Word and Object*. Cambridge, MA: MIT Press.

Reinhart, Tanya (1976). *The Syntactic Domain of Anaphora* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Reinhart, Tanya (2006). *Interface strategies*. Cambridge, MA: MIT Press.

Rice, Curt & Sylvia Blaho (2010). *Modeling Ungrammaticality in Optimality Theory*. London: Equinox.

Rizzi, Luigi (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.

Rosenbaum, Peter S. (1967). *The Grammar of English Predicate Complement Constructions.* Cambridge, MA: MIT Press.

Ross, John R. (1967). *Constraints on Variables in Syntax* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Rumelhart David. E., Geoff. E Hinton, Ronald J. Williams (1986). 'Learning internal representations by error propagation', in D. Rumelhart, J. McClelland, and the

PDP Research Group (eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1*. Cambridge, MA: MIT Press, 318–362.

Safir, Ken (2004). *The Syntax of Anaphora*. Oxford: Oxford University Press.

Samek-Lodovici, Vieri (2001). 'Crosslinguistic Typologies in Optimality Theory', in Geraldine Legendre, Jane Grimshaw, Sten Vikner (eds.), *Optimality-Theoretic Syntax*. Cambridge, MA: MIT Press, 315-354.

Samek-Lodovici, Vieri (2005). 'Prosody-syntax interactions in the expression of focus', *Natural Language and Linguistic Theory,* 23: 687-755.

Samek-Lodovici, Vieri (in press). *The Interaction of Focus and Giveness in Italian Clause Structure*. Oxford: Oxford University Press.

Samek-Lodovici, Vieri, and Alan Prince (2002). 'The fundamental properties of harmonic bounding', *Rutgers Optimality Archive*, ROA-785.

Smolensky, Paul and Géraldine Legendre (2006). *The Harmonic Mind: From Neural Computation to Optimality-theoretic Grammar*. Cambridge, MA: MIT Press.

Speas, Margaret (1997). 'Optimality theory and syntax: null pronouns and control', in D. Archangeli & D.T. Langendoen (eds.), *Optimality theory: An Overview*. Oxford: Blackwell-Wiley, 171-99.

de Swart, Henriëtte (2010). *Expression and interpretation of negation: an OT typology.* Dordrecht: Springer.

de Swart, Henriëtte and Ivan Sag (2002). 'Negation and negative concord in Romance', *Linguistics and Philosophy,* 25: 373-417.

Szendröi, Kriszta (2003). 'A stress-based approach to the syntax of Hungarian focus', *The Linguistic Review,* 20(1): 37-78.

Vikner, Sten (2001). 'The Interpretation of Object Shift and Optimality Theory', in G. Müller & W. Sternefeld (eds.), *Competition in Syntax*. Berlin: Mouton de Gruyter, 321-340.

Wexler, Kenneth (1994). 'Optional infinitives, head movement and the economy of derivations', in D. Lightfoot and H. Hornstein (eds.), *Verb movement*. Cambridge: Cambridge University Press, 305-362.

Wexler, Kenneth (1998). 'Very early parameter setting and the Unique Checking constraint: a new explanation of the Optional Infinitive stage', *Lingua*, 106: 23–79.

Zeevat, Henk (2001). 'The asymmetry of optimality theoretic syntax and semantics', *Journal of Semantics*, 17(3): 243-62.

Zeevat, Henk (2006). 'Freezing and Marking', *Linguistics,* 44(5): 1095-111.

Zubizarreta, Maria-Luisa (1998). *Prosody, Focus, and Word Order*. Cambridge, MA: MIT Press.

Zwarts, Joost (2004). 'Competition between word meanings: the polysemy of (a)round', in C. Meier and M. Weisgerber (eds.), *Proceedings of Sinn und Bedeutung 8*. Frankfurt/Konstanz: University of Konstanz, 349-60.

---

[i] For a discussion of the issue of ineffability see e.g. Pesetsky (1997) and Rice & Blaho (2010); for a critical treatment of functionally-driven OT see Newmeyer (2005). For more discussion of the nature of the challenges OT faces see Müller (2012).