Overall Description of the Research Unit and the Coordination Proposal

Cyclic Optimization

Jochen Trommer, Leipzig (spokesperson)

Cyclic structure building (as in the Minimalist Program, Chomsky 2007) and optimization (e.g., in Optimality Theory, Prince & Smolensky 2004) are two central research areas of grammatical theory, but only a fraction of the research in both fields addresses their interaction. The goal of the research unit (RU) proposed here is to explore the hypothesis that combining cyclicity and optimization achieves an added level of explanation not available to either alone: Cyclicity can substantially restrict the search space of optimization processes and explain why optimization is often surface-opaque, while optimization yields a principled account for the timing, size, and variability of cycles (McCarthy 2010, Trommer 2011,inkelas & Caballero 2013, Kiparsky 2015, Sande & Jenkins 2018, Müller 2020). Based on these results, we want to extend the empirical coverage of models where cyclicity and optimization crucially interact to a comprehensive range of grammatical phenomena, and develop their full theoretical potential.

<table>
<thead>
<tr>
<th>Project</th>
<th>Pls</th>
<th>Title</th>
<th>Institution</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mor Mor</td>
<td>Gereon Müller</td>
<td>Prospects of Inflectional Morphology in Harmonic Serialism</td>
<td>UL</td>
<td>Mor, Syn</td>
</tr>
<tr>
<td>Mor Phon</td>
<td>Jochen Trommer &amp; Morphological Strata of Tone Eva Zimmermann</td>
<td>Correlates of Affix Order</td>
<td>UL</td>
<td>Phon, Mor</td>
</tr>
<tr>
<td>Sem Phon</td>
<td>Barbara Stiebels &amp; Semantic and Phonological Jochen Trommer</td>
<td>Correlates of Affix Order</td>
<td>UL</td>
<td>Mor, Phon, Sem</td>
</tr>
<tr>
<td>Syn Syn</td>
<td>Fabian Heck</td>
<td>Syntactic Repairs and Cyclic Optimization</td>
<td>UL</td>
<td>Syn</td>
</tr>
<tr>
<td>Syn Phon</td>
<td>Philipp Weisser</td>
<td>Prosodically Determined Dislocation of Coordinators</td>
<td>UL</td>
<td>Syn, Phon</td>
</tr>
<tr>
<td>Syn Mor</td>
<td>Maria Kouneli</td>
<td>Layers of Morphosyntactic Number in Eastern Sudanic</td>
<td>UL</td>
<td>Mor, Syn</td>
</tr>
<tr>
<td>Syn Sem</td>
<td>Barbara Stiebels</td>
<td>Optimal Matches between Clause-embedding Predicates and their Clausal Complements</td>
<td>UL</td>
<td>Syn, Sem</td>
</tr>
<tr>
<td>Com Asp</td>
<td>Gregory Kobele</td>
<td>Computational Aspects of Cyclic Optimization</td>
<td>UL</td>
<td>Com</td>
</tr>
</tbody>
</table>

Table 1: List of projects (Syn:Syntax, Phon:Phonology, Sem:Semantics, Mor:Morphology, Com:Computational Linguistics), see section 2.2.1 for explanation of the code using the cycle (“Mor”) symbol; UL: Universität Leipzig

1 State of the art and preliminary work
In this section, we substantiate the claim that cyclicity (A) and optimization (B), although not always under this terminology, are ubiquitous in theoretical models for all areas of grammar, and sketch our view of the current situation in research combining cyclicity and optimization: a broad array of innovative developments, but still in need of focal points of convergence (C). Under (D), we summarize our own previous research related to this topic, documented by the publications in 1.1. In section 2, we will lay out how the proposed RU aims at a substantial contribution to the development of Cyclic Optimization.

A Cyclicity
The general intuition behind cyclic models of grammars is that the same operations (or types of operations) apply iteratively in increasingly bigger morphosyntactic domains. For concreteness, we will assume here a version along the lines of Kiparsky (1982), where cyclicity is implemented by interleaving syntactic or morphological structure-building operations with interpretative or restructuring grammatical processes (cf. Figure 1). Thus, in minimalist syntax (Chomsky 2001, 2007), the basic operation Merge is structure-building in creating hierarchical syntactic
representations (‘trees’) whereas Agree just modifies the featural content of syntactic units, or, under a slightly different perspective, adds information for morphological interpretation. For example, the derivation of the German sentence *Ich denke, sie glaubt* ‘I think she believes’ involves alternating applications of *n*erge and Agree, combining first *sie* and *glaubt* (by *n*erge), establishing agreement between them (→ *sie glaubt*) followed by further *n*erge applications for *ich* and *denk* and a final round of Agree matching these items (→ *ich denke, sie glaubt*). Cyclicity provides a natural explanation for the fact that verbs agree with the closest subject, hence that *Ich denkt, sie glaube* (where the outer verb matches the embedded subject and vice versa) is ungrammatical: At the point of derivation where Agree applies for the first time, the only noun present in the derivation is *sie*, and in the second Agree application, *glaub* is not longer accessible. Note also that cyclicity in general does not require strict (one-by-one) alternation between structure building and other operations. Thus, in more complex representations, several applications of *n*erge might be followed by several interpretative steps, still exhibiting the mutual ‘sandwiching’ for the two different types of operations shown in Figure 1.

The concept of cyclicity as a synchronic grammatical mechanism dates back to the beginning of modern grammatical theory (Chomsky et al. 1956, Chomsky 1965) (though its early formulation, the transformational cycle, did not assume strict interleaving, but application of operations to increasingly bigger portions of complete representations). Chomsky & Halle (1968) identified cyclicity as a central design feature of the phonology-morphosyntax interface. Thus, the regular stress pattern of English 5-syllable words with main stress on the penultimate and secondary stress on the first syllable (as in *ab. ra.ca.* *da.bra*) is overwritten in complex forms such as *E. *li. za. be.than*, with secondary stress on the second, not the first syllable, corresponding to the main stress of the base word *E. *li. za. be.th*. The classical cyclic account for this pattern (Siegel 1974, Bermúdez-Otero 2012) is interleaving morphological structure building and phonological stress assignment. Stress is first assigned to the root (*E. *li. za. beth → *E. *li. za. beth*) followed by suffixing -an (*E. *li. za. be.th → E. *li. za. beth-an*), and in turn a further cycle of stress assignment establishing a new primary stress on the penultimate syllable and degrading the inherited primary stress to secondary status (*E. *li. za. beth-an → E. *li. za. be.th*).

Cyclicity is essential for virtually all areas of grammar in their interaction with syntactic or morphological structure. In compositional approaches to meaning, semantic interpretation is interleaved with morphological and syntactic composition (Kobele 2006), and conversely cyclicity also governs partial non-compositionality. Consider for example the derivation *be* ‘CAUS’ + *Staub* ‘dust’ → [be-[stäub]n]v in German, a verb with the idiosyncratic meaning ‘inpollinate’. If this undergoes further derivation, its idiosyncratic meaning is retained as in [be-[stäub]n]v + ung → [be-[stäub]n]v-ung ‘inpollination’. In word syntax approaches, this falls out from the fact that assignment of compositional (*bestäub → bestäubung*) and of non-compositional meaning (*Staub → be-stäub*) apply cyclically, i.e., after corresponding affixation steps, based on the meaning established in the previous cycle (see Bermúdez-Otero 2012 on a proposal integrating lexicalization in a cyclic architecture). Even in the interaction of syntax and prosody, which has long been claimed to work non-cyclically (Kiparsky 1982, Nespor & Vogel 1986), there is growing evidence of cyclicity (see Wagner 2010, McPherson & Heath 2016, Wierzba 2017, and Scheer 2009 for critical discussion).

Cyclicity provides a direct explanation for one of the central and most pervasive empirical generalizations on linguistic systems: The properties of bigger grammatical units depend asymmetrically and locally on properties of smaller units (Chomsky 2012:1). Thus, *Bestäubung* locally inherits the meaning of *bestäuben*, but not of deeper embedded *Staub*, and neither does *Staub* itself inherit the meaning of *bestäuben*. Similarly, whereas bottom-up cyclicity abounds in stress systems crosslinguistically (see, e.g., Halle & Vergnaud 1987), anticyclic stress systems are apparently undocumented (see Steriade 2008 for one of the few potential cases and Bermúdez-Otero 2018a for a reassessment).

Whereas cyclicity is a standard assumption in theoretical grammar, there have been attempts to eliminate it. Thus, it was a major tenet of Government and Binding Theory in the 1980’s that all structure-building operations precede all other operations, giving up the cyclic elements in previous versions of transformational grammar, a move revoked in Chomsky’s minimalist papers (Chomsky 2001, 2007) based on the insight that the non-cyclic architecture does not allow for a principled account of locality and economy effects. Similarly, early work in parallel optimality-theoretic phonology aimed to eliminate
the cyclic architecture of Lexical Phonology in favor of parallel computation, based either on prosodic structure or on paradigmatic relations (Benua 1997, Burzio 2005, McCarthy 2005). The clearest evidence for cyclic phenomena which cannot be captured by sensitivity to prosody are tonal opacity effects in syntax, as in Mandarin Chinese 3rd tone sandhi which opaquely reflects syntactic cycles (Kaiser 1985, see also Jones 2014 on Kinande). This also provides direct evidence against the claim of non-cyclicity specific to the syntax-phonology interface (Nespov & Vogel 1986, Selkirk 1995).

Capturing prima-facie cases of cyclic opacity by the appeal to paradigmatic relations has also encountered difficulties due to the lack of output bases (Bobaljik 2008, Trommer 2013, Mascarel 2016). Symptomatically, a major result of work on opacity effects in OT was that the most successful implementations involve asymmetric relations between members of a morphological paradigm, such as Elizabeth and Elizabethan, where the simpler form unidirectionally determines phonological properties of the more complex form (Benua 1997, Hall & Scott 2007). Thus, as shown by Bermúdez-Otero (2007), while the paradigmatic cycle in Benua makes slightly different empirical predictions than Lexical Phonology, it effectively is a cyclic formalism where derivational steps are supplanted by recursion (see below and ComASP for more discussion of non-derivational interpretations of cyclicity).

Whereas the most straightforward understanding of cyclicity is in derivational terms as a successive alternating application of grammatical operations of different type, there are influential alternative interpretations. Thus, the original conception of stress cycles in Chomsky & Halle (1968) was built on a complete contiguous process of syntactic operations building, preceding all phonological processes. In this system, the phonological cycle then recapitulates the original structure building process by the alternating application of phonological rules and the erasure of inner brackets that function as temporary traces of the applications of phrase structure rules. Thus, for Elisabethan, the syntax would first generate [[Elisabeth] an] (still without any stress), then perform an initial cycle of stress assignment to the innermost bracket-less subconstituent Elisabeth, resulting in [[E.‘.li.sa.beth]an], followed by a first round of bracket erasure, leading to E.‘.li.sa.be.than], which the second cycle of stress rules map to E.‘.li.za.‘be.than] by reassigning a new primary stress, and preserving the stress of the first cycle as secondary stress. This ‘recapitulative’ interpretation of cyclicity makes the same predictions in many cases, and whereas standard Lexical Phonology (Kiparsky 1982) has abandoned it in favor of the conceptually simpler direct-interleaving approach, which obviates the use of boundary symbols in phonological computation, it has remained an influential alternative (Halle & Vergnaud 1987). For empirical arguments favoring an interactionist (non-recapitulative) interpretation of cyclicity, see Bresnan (1971), Legate (2003) on evidence from sentential stress, Dloatian (2020) for phonologically conditioned suppletion, and Scheer (2011) on a more general discussion of interactionism. A hybrid version of recapitulative cyclicity is still a standard assumption in Distributed-Morphology approaches to morphological spillout, where abstract (phonologically unspecified) clause- (or phase-)sized syntactic structure building is followed up by a second round of structure building – Vocabulary Insertion – which iteratively realizes heads in word-sized constituents inside-out (Trommer 1999, Bobaljik 2000, Embick 2010) by morphological VIs with phonological content recapitulating the original syntactic structure building.

A non-derivational interpretation of the cycle without reference to paradigms has been developed in sign-based phonology by Orgun (1996), who shows that standard versions of cyclic architecture can be equivalently recast in a declarative formalism (see also Stump 2001 for a declarative morphological formalism with basically cyclic design features).

Surprisingly, the mathematical nature of cyclicity and the formal relation between its different implementations and with locality principles in non-cyclic approaches such as HPSG and Sign-Based Construction Grammar (see Sag 2007, Müller 2016 for discussion) have hardly been addressed in the literature. A central goal of project ComASP is to lay mathematical foundations for these questions.

⑤ Optimization
By optimization, we understand a grammatical procedure which chooses among different potential alternatives and blocks suboptimal alternatives, based on a structured set of criteria. A classical case is the Elsewhere Condition (Kiparsky 1973), which captures the fact that more specific morphological realizations block more general ones (see Williams 2007, Fanselow 1991 for extensions to syntax). Thus, the plural allomorphy in cases such as box ~ box-es/*box-en vs. ox ~ *ox-es/ox-en can be
captured by two conflicting desiderata: ① Realize a nominal plural by -s and ② Realize the nominal plural of ox by -en, where both operations are in principle applicable to ox, and the Elsewhere Principle resolves to choose ② based on its greater specificity. Whereas there is broad consensus that the Elsewhere Principle governs morphological systems (Halle & Marantz 1993, Stump 2001, Brown & Hippenstey 2012), it is often not appreciated that it is a complex and controversial mechanism. Thus, as shown by Müller (2005), standard versions of the principle invoke at least three different criteria (1) specificity by context sensitivity (e.g., to specific lexical items), (2) specificity by the number of specified substantial features, and (3) qualitative asymmetries between different features (Noyer 1992).

Optimization in the broad sense envisaged here underlies psycholinguistic race models (Hay 2000) and analogy-based approaches to morphology and phonology (Albright & Hayes 2003, Arndt-Lappe 2014), but also the notion of last resort and repairs in derivational approaches to syntax (Chomsky 1991, Bjorkman 2011), where operations such as do-insertion and the use of resumptive pronouns have been argued to represent last-resort options to repair otherwise ungrammatical representations, and more generally any kind of transderivational constraints where grammaticality is computed by comparing different related derivations in parallel (see, e.g., Graf 2013). A maximally general optimization formalism is Harmonic Grammar (Smolensky & Legendre 2006, Potts et al. 2010), which is based on numerically weighted constraints, allowing for decision processes by arithmetic summing up of constraint violations, and the most successful optimization approach developed so far is Optimality Theory (Prince & Smolensky 2004), which can be interpreted as a more restrictive version of Harmonic Grammar, where constraints have to be strictly ranked. The major motivation for the success of OT in the phonological research of the last decades have been cases of rule conspiracies where different operations compete (or cooperate) to derive the same result. Thus, in Mokilese (see Paschen 2018 and references therein), the progressive is formed by copying (reduplicating) the first syllable of the base in a way which renders it phonologically heavy. If the base syllable contains a consonant after its first vowel, heaviness is achieved by copying this consonant along with the preceding vowel (\textit{podok} ‘to plant’ $\rightarrow$ \textit{pod-podok} ‘planting’), but if there is no postvocalic consonant, the reduplicative syllable is rendered heavy by lengthening the copied vowel (\textit{via} ‘to do’ $\rightarrow$ \textit{wi-wia} ‘doing’). This pattern can be captured in OT by a constraint requiring a heavy-syllable reduplicant (RED=$\alpha_{\text{h}}$) outranking a constraint against long vowels ($^{*}$V:) outranking in turn a constraint against segment copying ($^{*}$COPY, see Figure 2).

A seminal result of OT-phonology is that constraint interaction captures not only blocking of processes by iso-functional processes (such as blocking of lengthening by copying in Mokilese), but also by functionally independent processes. Thus, in many languages, lexical high tones are deleted if they immediately follow other high tones, a pattern commonly attributed to a constraint against adjacent identical tones called OBLIGATORY CONTOUR PRINCIPLE (OCP). As shown by Myers (1997), in Shona, the OCP blocks an independent process of high tone spreading, a pattern which cannot be directly captured in a purely derivational account, but follows directly in OT by ranking of the OCP above the constraint(s) triggering tone spreading.

OT has also opened up a new perspective on language-specific variation and parametrization, reducing differences between languages to different repair options triggered by distinctive constraint rankings, resulting in a crosslinguistic version of the conspiracy pattern. Thus, as shown by Pater (1999), the ban on voiced obstruents after nasals is effected by different alternations such as assimilation, consonant deletion, or coalescence. The same abstract pattern also characterizes many morphosyntactic processes, see, e.g., Grimshaw (2001), Trommer (2006, 2008), Nevins & Sandal (2011), Gerlach (2001) on morpheme cooccurrence restrictions, Trommer (2003c, 2008) on affix linearization, Wunderlich (2000) on morphological ‘taboos’, Aissen (2003), Wunderlich (2008), Stiebels (2000, 2002), Woolford (2001), de Swart (2007), Keine & Müller (2015) on case, and Legendre et al. (1998), Ackema

<table>
<thead>
<tr>
<th>Base: podok</th>
<th>RED=$\alpha_{\text{h}}$</th>
<th>V:</th>
<th>COPY</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. p$\cdot$pc.$\cdot$dok</td>
<td>$^{*}$!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. p$\cdot$pc.$\cdot$dok</td>
<td>$^{*}$!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. pc.$\cdot$pc.$\cdot$dok</td>
<td></td>
<td></td>
<td>***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Base: wia</th>
<th>RED=$\alpha_{\text{h}}$</th>
<th>V:</th>
<th>COPY</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. wi.wi.a</td>
<td>$^{*}$!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. wi.wi.a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. wi.wi.a</td>
<td>$^{*}$!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Mokilese Reduplication

Crucially, OT instantiates a type of optimization more powerful, but otherwise isomorphic to last-resort repairs in syntax and the Elsewhere Principle: a selection process between alternatives (e.g., default case vs. assigned case, different allomorphs or phonological forms) based on grammatical criteria (e.g., the OCP vs. the preference for default case), and blocking of alternative options (DPs do not exhibit default case in addition to assigned case). Systematic blocking is what establishes optimization as a disjunctive procedure, i.e., an operation which cannot be reduced to a strictly ordered sequence of sub-operations. Thus, Baković (2011) shows that morphological elsewhere blocking is not amenable to a purely derivational analysis, and the same seems to hold for last-resort operations in syntax.

Cyclic Optimization

By Cyclic Optimization (CO), we understand systems which invoke both iterative optimization and iterative morphosyntactic structure building such that cyclicity emerges from the interaction of both mechanisms. CO approaches originate in three different lines of research: First, approaches combining standard OT-optimization with elements of rule-based Lexical Phonology assuming qualitatively different phonological subgrammars that are cyclically interleaved with morphosyntactic levels of increasing size such as Stem, Word and Phrase (Kiparsky 2000, 2015, Bermúdez-Otero 2012, Rubach 2003). In fact, most work in OT has at least implicitly subscribed to a minimal stratal architecture of this type with respect to the interaction of lexical and phrasal phonology (McCarthy & Prince 1993, Myers 1997). The major motivation for this line of research has been the problem strictly parallel OT faces with opacity. However, foundational work in Stratal OT has shown that CO approaches derive further important empirical predictions, such as a natural upper bound on opacity (Bermúdez-Otero 2003), bracket erasure (the unavailability of morphological information at a stratum at subsequent strata, Bermúdez-Otero 2012), or the ‘life cycle’ of phonological processes (Ramsamy 2015). CO also underlies the major formalizations of Construction Phonology (Orgun 1996, Inkelas 2014).

Second, a version of CO has independently been developed in derivational approaches to cyclic effects in syntax departing from a major controversy in minimalist syntax whether at a given stage of derivation grammars uniformly favor Nørge or Agrecht operations. As shown by Müller (2009), assuming that languages parametrically choose either preference accounts for the difference between ergative-absolutive and nominative-accusative languages, and can be naturally captured by violable constraints. An analogous type of competition holds in cyclic accounts of languages where verbs alternate between subject and object agreement. A standard assumption here is that the agreement probe at a specific point differentially ‘prefers’ one accessible argument over the other, but turns to the dispreferred argument if agreement with the preferred one fails, and resorts as last resort to ‘default agreement’ (Béjar & Režič 2009). Again, this provides a case where derivational cyclicality cannot replace, but requires an optimizing selection process among alternative choices. Finally, there is growing evidence that also the third central mechanism of derivational minimalism, viz. establishing a phase, does not apply deterministically in derivations, but can be delayed, resulting in ‘dynamic’ phases (Bobaljik & Wurmbrand 2005, den Dikken 2007, Bošković 2014, Harwood 2015) or apply only partially to subparts of the processed representation (D’Alessandro & Scheer 2015, Newell 2018).

The third and most recent line of research leading to CO is work in iterative optimization approaches, especially Harmonic Serialism (HS) and Optimality Theory with Candidate Chains (OT-CC), which have also been motivated by the problems of parallel OT with phonological opacity (McCarthy 2007) and the ‘Too-Many-Solutions Problem’, the problem that non-iterative versions of OT predict unattested repair mechanisms (McCarthy 2008a, 2010). HS differs in two respects from fully parallel OT: First, optimization steps can only make single changes to a given form, considerably restricting competition to a finite candidate set, and second, optimization is iterative. As long as there is an output candidate which further optimizes the output of the last optimization step, optimization is reapplied. In OT with Candidate Chains (OT-CC, McCarthy 2007), a variant of HS generates different chains of serial optimization, which are then globally evaluated by an OT-grammar including constraints that restrict the order of optimization steps. While the computational complexity of OT-CC has been justified with its merits for phonological opacity, HS can account for specific cases of opacity (see, e.g., McCarthy 2008b), and in the last years different conservative extensions to HS have been developed which substantially increase the ability of
the theory to cover opacity (e.g., Jarosz 2014).

Even in their purely phonological applications, HS and OT-CC show a type of proto-cyclicity. Operations of one type \( T_1 \) (e.g., syllabification) may be followed by operations of a second type \( T_2 \) (say foot formation) followed again by a type \( T_1 \) process, though this does not involve morphosyntactic structure building. Under the natural assumption that a single serial OT-grammar optimizes both morphology and phonology choosing between operations from both based on ranked constraints, HS inherently instantiates a full-fledged cyclic model of grammar. A version of OT-CC which makes this explicit is the Optimal-Interleaving (OI) framework by Wolf (2016) where spellout and phonological processes are competing operations on a par selected by iterative OT evaluations (see Wolf 2015 for a HS-version).

Whereas CO has clear roots in specific research traditions, it is in principle compatible with a wide variety of assumptions on representation, computation, and grammatical architecture: symbolic or gradient representations (see, e.g., Plag et al. 2017 on phonetic gradience in the representation of morphemes, and Mor Phon for phonological gradience in a CO setting), exemplar-based approaches (see, e.g., Inkelas 2016), postsyntactic or lexicalist architecture (SynMor vs. Mor Mor).

(2) Preliminary Work by the Applying Researchers

The RU brings together three threads of our research:

First, the Empirical Domain of Optimization Processes: Leipzig is arguably one of the central European research institutions for theoretical work on Optimality Theory; the work by the applicants has been instrumental in extending the coverage of optimization-based explanation to substantially new empirical areas. Thus, extensive work by Heck and Müller explores optimization in phrasal syntax (Heck & Müller 2007, 2013, 2016, Müller 2000a, 2003, 2009, 2015, and the project DFG MU 1444/2-1,2 Optimalitytheoretische Syntax des Deutschen). Heck (2001) (together with Wilson 2001) presented the first approach employing CO in syntax. Müller (2000b) is the standard reference work on optimality-theoretic syntax. Foundational work by Stiebels (1999, 2000, 2006) has established optimization for the interaction of lexical semantics and argument linking. Trommer (2001a) provides the first comprehensive extension of optimization to postsyntactic morphology, further developed in work on affix order (Trommer 2001b, 2003c, 2008), consequences for grammatical architecture (Trommer 2002a) and hierarchy effects (Trommer 2003a, 2006, 2008). Research by Kobele has been essential in establishing the mathematical complexity of OT (Heinz et al. 2009). Our work on other optimization methods covers the Elsewhere Principle (see, e.g., Müller 2005 for formal elaboration, Trommer 2015b for a reformulation as licensing conditions, and Weisser 2007 for an alternative, the Nearest Neighbor Principle). Our recent research on constraint-based optimization has focused on variants of optimization departing from standard OT to further extend its empirical coverage. Thus, joint work by Trommer and Zimmermann has established the potential of Containment Theory (van Oostendorp 2008) for nonconcatenative morphology (Zimmermann 2013c, 2014, 2017, Trommer 2011, 2014, 2015a, Zimmermann & Trommer 2014, 2015) and phonological opacity effects (Trommer 2016c, Zimmermann & Trommer 2016, Trommer 2016a, 2017). Our recent research in syntax and morphology proper establishes a convergence with phonology-born HS (McCarthy 2010) (see Heck & Müller 2016 on HS in syntax and Müller 2020 for inflectional morphology). Zimmermann (2019a) applies HS to lexical conditioning of accent systems, and Trommer & Bank (2017) to morphological learning. Our work in HS is a primary point of departure for the proposed research group (especially MorMor and SynPhon), as is recent work on integrating Harmonic Grammar and Gradient Symbolic Representations to capture scalar and cumulative effects in morphophonology (Trommer 2016b, 2019, Zimmermann 2018a,b, 2019b).

Second, Monotonicity of Basic Operations: A unifying desideratum in grammatical research of the last decades is the assumption that grammatical systems work in a strictly monotonic way, typically by requiring that all operations in a grammatical component are strictly information-adding (e.g., the Inclusiveness Condition of Chomsky 2000). Standard cyclic architectures such as Bare Phrase Structure and Lexical Phonology are classical cases of monotonic systems. Our work on monotonicity has focused on the controversial question whether different parts of grammar are monotonic in the same or in fundamentally different ways. Thus, research by Trommer and Zimmermann explores the assumption that phonological computation does allow addition, but not removal of material (Containment Theory) (Zimmermann 2013c, 2014, 2017, Trommer 2011, 2014, 2015a, 2016c,a, 2017, Zimmermann & Trom-
mer 2014, 2015, 2016, Zimmermann & Topintzi 2014), just as typically assumed for syntax. Kobele (2006, 2018, 2012) pushes monotonic syntax-semantic mapping to its logical limit. While Alexiadou & Müller (2008) extend a purely structure-building approach to inflectional morphology, Trommer (2001a, 2003b) and Henze & Zimmermann (2010, 2011) argue that morphological spellout is monotonically consumptive. An important point of departure for the RU is the hypothesis that monotonocity is not strictly linked to structure building or removal, but to sets of abstract features triggering grammatical operations which are monotonically discharged, an assumption spelled out in the DFG Koselleck project Syntaktischer Strukturabbau (Müller 2017, 2018, see also Heck 2016).

**Third, Grammatical Interfaces of Morphology and Syntax:** The interfaces between different grammatical modules have played a prominent role in the research of all applicants. Kobele (2006, 2012, 2018) provides crucial contributions to a minimalist model of the syntax-semantics interface, and Stiebels (2000) for morphology and semantics. Evidence for a direct interaction between syntax and morphology is given by Trommer (2001a, 2002b), Heck et al. (2009), Keine & Müller (2015), Guseva & Weisser (2018), Weisser (2019), and Kouneli (2019). The nature of the morphology-phonology interface has been the central issue in the DFG-Network *Core Mechanisms of Morphological Exponence* (project TR 521/2-1). Recent work by Trommer and Zimmermann (Zimmermann 2013c,b, 2014, 2015, Trommer & Zimmermann 2014, Zimmermann & Trommer 2015) has shown that many classical problems of non-concatenative morphology can be solved by a refined notion of prosodic affixation (see also Müller 2002, 2013, for an approach to morphology where basically all exponence is phonological), an approach extended to a broader range of morphophonology in the recently completed DFG-project on *Featural Affixes* (Jochen Trommer). This line of research has developed a unified approach which covers not only non-concatenative morphology, but also morphologically conditioned phonology and forms a substantial part of the new Emmy-Noether research group on *Grammatical Strength in Prosodic Morphology* (Eva Zimmermann) and provides essential impulses or the phonology-related projects of the RU (e.g., Mor\*Phon and Sem\*Phon). Our expectation is that the RU not only allows for a substantially new take on interfaces (see, e.g., the innovative models assumed in Mor\*Phon and Syn\*Mor), but also for comparing closely related empirical phenomena such as affix order facts under different grammatical architectures (e.g., under a lexicalist perspective in Sem\*Phon and Mor\*Mor, and under a post syntactic morphology approach in Syn\*Phon and Syn\*Mor), and to identify formal parallels across different interfaces (Com\*Asp).

The RU also develops impulses of the Leipzig DFG research training group on *Interaction of Grammatical Building Blocks* (IGRA) (GRK-2011) (see, e.g., the dissertations of Murphy 2017, Zaleska 2018, Kushnir 2018, Hein 2018, Paschen 2018, Gjersoe 2019, with contributions to CO). The proposed RU is in line with IGRA in addressing not so much the interaction of grammar with extra-grammatical factors, but in and between grammatical modules; but, where IGRA highlights parallels in the interaction of all grammatical operations and constraints, the RU proposal focuses on two of the central instances of ‘building blocks’: derivational structure building and optimization, with a central role assigned to the interfaces of grammatical modules (see above). A thematic extension of IGRA’s focus on phonology, morphology, and syntax is the inclusion of semantic aspects of cyclicity.

### 1.1 Project-related publications by members of the research unit


2 Objectives and joint work program

2.1 Objectives of the overall project and expected benefits of collaboration within the unit, incl. a description of the group composition and their project-specific qualifications

The ultimate goal of theoretical linguistics is to identify general cognitive principles underlying the empirical abundance of natural language; cyclicity, optimization, and their combination are prime candidates for such principles. Against this background, the constitution of the RU conveys three major benefits, which could not be achieved by the single projects alone: First, it allows for systematically investigating CO across grammatical subfields (see section 2.2.1 and Com:Asp) by fostering the collaboration of researchers with primary specializations in all of the involved areas, phonology (especially Trommer and Zimmermann), morphology (especially Müller, Stiebels, and Trommer) syntax (Heck, Weiss, Kounelii), and semantics (Stiebels, Kobele, see also the complete list in Table 2), to identify abstract mechanisms which are independent from specific grammatical modalities. Second, the RU allows for investigating different competing versions of CO in parallel in overlapping and closely related empirical domains (see section 2.2.1 A) for details). This guarantees a stringent evaluation of diverging hypotheses, but also provides an essential environment for theoretical cross-fertilization among different approaches (see section 2.2.1 C). Third, the RU allows for covering the full span of relevant research methodologies for a better understanding of general grammatical principles, ranging from the central mathematical Com:Asp project, which addresses abstract properties of CO across different implementations, over a group of projects focusing on the development and evaluation of specific theoretical approaches (Mor:Phon, Mor:Mor, and Syn:Syn) to projects with a crucial empirical component in fieldwork (Syn:Mor), crosslinguistic typology, and the work with native-language informants (Sem:Phon and Syn:Phon) to make crucial new data for the evaluation of CO accounts accessible. Again, the range of expertise and scientific man-power to cover this broad spectrum goes far beyond the capacity of single research projects. Table 2 lists all the applicants with projects and primary research areas. Apart from the fact that the composition of the group covers specializations in all major sub-areas of grammar, and multiple methodologies (see above), the participating researchers also have a strong background in research on cyclicity (Kouneli and Weiss), optimization (Zimmermann and Stiebels) or both (Heck, Müller, Kobele, Trommer). See section 1 A, B, C for a detailed record.
The RU pursues two specific goals which we regard as essential next steps for evaluating the merits of CO for the theoretical modeling of language: First, to explore empirical and formal properties of CO systems. Second, to develop the potential of CO to solve central problems for purely cyclic and purely optimizing approaches. In what follows, we spell out how the proposed RU addresses these goals.

2.2 Joint work program including proposed research methods

2.2.1 Exploring formal and empirical properties of CO systems

**A) Extending the Empirical Coverage of CO:** The viability of CO has so far only been tested on a restricted set of phenomena, and with a bias to few languages, even in the interaction of morphology and phonology. A major goal of the RU is therefore to systematically explore the empirical consequences of this approach in all areas of grammar which show a potential for cyclic interaction, but also its potential limitations (see especially Syn\(\rightarrow\)Sem and Syn\(\rightarrow\)Syn).

Under the standard assumption that grammatical architecture comprises the four modules Phonology, Morphology, Syntax, and Semantics, and that structure-building operations are limited to Morphology and Syntax, cyclic systems may involve all and only the mutual interactions of Morphology and Syntax with all other grammatical modules (including cyclic interaction between Morphology and Syntax, and intra-modular interactions in Morphology and Syntax), as shown in Figure 3, where \textbf{X:Y} (in the text and in project names: \textbf{X} \(\rightarrow\) \textbf{Y}) denotes cycles with structure building operations in module \textbf{X} and interleaved operations from module \textbf{Y}.

![Figure 3: Possible cyclic interactions](image-url)

All these interactions are covered by single projects in the RU (see the list of projects in 1.2). Many of them have not or only marginally been addressed under the perspective of CO. Thus so far there exist only sporadic pioneering applications of CO in morphology proper (Mor\(\rightarrow\)Mor), and for the interaction of syntax with morphology (Syn\(\rightarrow\)Mor) and semantics (Syn\(\rightarrow\)Sem, but see de Hoop 2013). More specific empirical gaps for CO are filled for the modular interactions which have been crucial to the development of CO. Thus, Mor\(\rightarrow\)Phon tackles tonal phonology which has been almost completely neglected in the CO literature (but see Jones 2014) although it arguably provides some of the most
tricate data suggesting cyclicity. Sem\textsuperscript{Mor}Phon addresses affix order which has been of central importance in Lexical Phonology and Morphology, but has not been addressed under a CO approach (note that we use Sem\textsuperscript{Mor}Phon as a shortcut for Mor\textsuperscript{Mor}Sem+Mor\textsuperscript{Mor}Phon, and that ‘Com\textsuperscript{Mor}Asp’ is a simple abbreviation since the project covers abstract properties of all grammatical domains involved here).

\textbf{B} **Comparative Evaluation of Different Co Approaches:** A major synergetic effect we expect from the RU is to elaborate differential properties of CO approaches by applying different versions to closely related empirical phenomena. Thus, the projects employ, in a balanced way, the major existing approaches to CO (albeit often in combination and in innovative variations), as shown in Figure 4. At the same time, there is a deliberate overlap in the type of empirical phenomena which are prominent in different projects. Thus, linearization of affixes (and clitics) is highly prominent in Mor\textsuperscript{Mor}Mor, Syn\textsuperscript{Mor}Phon, and Syn\textsuperscript{Mor}Mor. Mor\textsuperscript{Mor}Phon and Syn\textsuperscript{Mor}Phon both raise the issue of boundary phenomena between phrasal and lexical phonology, and multiple exponent is central in both Mor\textsuperscript{Mor}Mor and Syn\textsuperscript{Mor}Mor. We expect that workshops, and teams addressing these phenomena across sub-projects will elucidate both the empirical similarities and convergences between different CO formalisms. In addition, Com\textsuperscript{Mor}Asp addresses differences and equivalences between different CO versions from the more general and abstract perspective of computational complexity theory.

\textbf{C} **Exploring Formal Properties of CO Systems:** Existing cyclic theories are conglomerates, where different types of cyclicity interact with inherently non-cyclic components of the grammar. They are framed in homogeneous theoretical vocabulary, obscuring shared functionality. These factors multiply if cyclicity is combined with different versions of optimization so that it remains largely unclear what the specific empirical predictions of specific CO approaches are, and why they differ empirically. Therefore a central goal of the RU is the systematic exploration of theoretical properties of existing CO systems.

CO systems may be classified into two major types, those with optimizing superstructure: where iterative optimization determines cycles (or avoidance of cycles), as in OI (Wolf 2016), and those with a structure-building superstructure, where optimization implements derivational steps in an independently timed (not-optimization-based) cyclic system (e.g., Stratal OT, Bermúdez-Otero 2018b, and Construction Phonology, Orgun 1996). Moreover, CO approaches differ wrt. a number of design features, which are at least partially orthogonal. First, both optimization and structure-building might differ in the degree of their internal iterativity. They can work strictly iteratively (as in HS for optimization) or in a batch mode as in Stratal OT, where typically a single phonological optimization step follows a sequence of affixation steps. In many syntactic CO systems, optimization applies only for every clause (Ackema & Neeleman 1998), phase (Fanselow & Čávar 2001, Müller 2000a) or phrase (Müller 2000c, Fischer 2004, Heck 2004), i.e., in batch mode, whereas Heck & Müller (2013) argue for strictly iterative optimization. Second, CO systems might interact crucially with layering, i.e., the assumption that recursive structure building inherently implies a fixed sequence of constituents. Thus, in Stratal OT, the assumption of at least minimal morphosyntactic layering (Stem > Word > Phrase) and basing the cyclic superstructure of phonological cycles on this system, leads to the consequence that a word such as yes will pass through all three cycles, even if it is not combined with other overt morphosyntactic material. On the other hand, most work on CO couched in OI explicitly rejects layering (Wolf 2016, Müller 2020). Third, a further orthogonal difference is whether cycles allow qualitative differences, i.e., different constraint rankings for different cycles, again dividing Stratal OT (restricted qualitative differences) and loop approaches (HS and OI) which presuppose globally consistent OT-grammars. Considering interfaces, cyclic models also differ in their modularity. Thus, Wolf (2016) argues for a single constraint ranking which governs all of morphology, syntax and phonology. Bermúdez-Otero (2007) implicitly assumes that there are two completely independent optimization procedures. Finally, CO systems employ different Optimization Formalisms, such as OT, HG, transderivational constraints, or Elsewhere block-

<table>
<thead>
<tr>
<th>Stratal or Layered OT/HG</th>
<th>Sem\textsuperscript{Mor}Phon, Mor\textsuperscript{Mor}Phon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase-based</td>
<td>Syn\textsuperscript{Mor}Phon, Syn\textsuperscript{Mor}Mor</td>
</tr>
<tr>
<td>Harmonic Serialism</td>
<td>Mor\textsuperscript{Mor}Mor Syn\textsuperscript{Syn}</td>
</tr>
<tr>
<td>Repair Rules</td>
<td>Syn\textsuperscript{Syn}, Syn\textsuperscript{Mor}Sem</td>
</tr>
</tbody>
</table>

Figure 4: CO Approaches in Different Sub-Projects
ing. Three questions arise from this parameter space: ① What is the complexity of these systems and how does complexity relate to the differential features of these formalisms? ② Where do these systems differ empirically in the patterns they can generate? ③ How would different combinations of these parameters work? The computationally oriented project ComASP which is complementary to the more empirically oriented projects will centrally address question ① and at an abstract level question ②. The latter will be a joint effect of different sub-projects working on related phenomena (see above). Question ③ will be addressed in detail in the second period of the RU, but is already on the agenda of single projects in this application such as MorPhon, which attempts an innovative combination of layering with HG.

### 2.2.2 Addressing central problems for optimization and cyclicity

Cyclic Optimization describes not a single monolithic framework, but a family of formalisms, each with specific empirical predictions. These are evaluated in detail in the single sub-projects, also in comparison to radically different theoretical approaches (see, e.g., SemPhon, MorMor, and SynSem). To estimate the merits of CO at a more abstract general level, a central strategy of the RU is to contrast it with approaches which use only cyclicity or only optimization. More specifically, our goal is the employment of CO to solve long-standing problems for purely cyclic and purely optimizing models. Since a significant amount of current work in theoretical grammar makes critical use of one of these mechanisms, these results would be relevant throughout the field. Here, we discuss four representative examples: ④ Locality and Myopia (problematic for purely optimizing approaches), ⑤ Cyclic Timing, ⑥ Lenient Cycle Effects (problematic for purely cyclic accounts), and ⑦ Heterogeneity of Cycles ≈ Construction-specific Optimization (problematic for both):

④ **Locality and Myopia**: A well-known problem for optimization formalisms is that processes appear to be optimizing, but only at an intermediate level. They are derivationally myopic (Collins 1997, Kimper 2012). Thus, in Kinande sentences, High tones generally shift one syllable to the left, which obscures previous optimization for tonal identity in morphological noun reduplication (Jones 2014). Myopia is also central in the cyclic derivation of the ungrammaticality for *[There seems someone to be in the room]* in Chomsky (1995, 2000), which is blocked because at the point of evaluating to be someone in the room the ranking of Merge > Move enforces merging of there → there to be someone in the room which by further Merge of seems and Move of there results in correct *[There seems to be someone in the room]*. Crucially, without cyclic myopia, both orders should be well-formed since both involve the same number of steps. Theories deriving myopic effects of course face the complementary challenge of phenomena that seem to be more global than single cycles. Thus, as shown by Inkelas (2016), relative restrictions on the order of affixes cannot be implemented in a model where every affixation step is an independent cycle. Also, phenomena such as repairs of movement out of island violations by resumptive pronouns (e.g., *the man who(m) I don’t believe the claim that anyone saw him*) seem to instantiate ‘supercyclic’ globality since the resumptive pronoun is merged in an inner cycle as a response to a constraint violation in an outer cycle (Müller 2015:38).

The innovative goal of the RU is to explore how cyclic optimization can capture (non-)myopia in morpho-syntax. We want to build here on first promising results in the modeling of extended experience (Müller 2020, MorMor, SynMor), but myopia is also relevant for repair phenomena as German expletive insertion (Heck & Müller 2013, SynSyn), in phonologically conditioned affix order (SemPhon), and in cases of recursive cleftization (SynPhon). ComASP investigates the hypothesis that specific types of restricted myopia in CO systems do not increase formal complexity.

⑤ **Cyclic Timing**: A central problem for derivational cyclic models is to provide a principled account of what triggers cycles at a specific point in a derivation, i.e., what governs the timing of structure-building and interpretive operations. As outlined above, this has been one of the main incentives to develop CO approaches in syntax, and the substantial empirical goal of the RU is to extend the coverage of these approaches in syntax (SynMor), e.g., to ‘dynamic phases’ (Bošković 2014, Wurmbbrand 2017), but also in morphology (MorMor) and semantics (SynSem). The theoretical goal behind this enterprise is to replace diacritics used in derivational theories as cyclic triggers by substantial constraints. Thus, the ‘strength’ of features triggering early or overt movement in minimalist syntax is a
purely arbitrary stipulation, but can can be reinterpreted in a CO system via higher ranked constraints requiring movement (affecting earlier cycles). First steps towards the elimination of morphophonological triggers have already been made by Bermúdez-Otero (2012), who argues that whether stem-level constructions show level-internal cyclicity is not due to stipulation, but to competition of stored forms with analytic morphology, reflected by the impact of frequency data on cycle effects (see Zuraw 2000 for a related approach). Optimization also promises to shed light on two other problems for a diacritic account, the dependence of cyclic behavior not only on affixes, but also on bases (Giegerich 1999), and the fact that cyclicity is often not morpheme-specific, but a feature of morpheme class or function.

**C** **Lenient Cycle Effects**: A major unexplored consequence of CO at the heart of the RU is the possibility to reinterpret apparently hard-wired cyclic opacity effects as emergent properties of viable constraints. Thus, the Bracket Erasure Convention (Pesetsky 1979) requires that morphological brackets are deleted after every cycle (e.g., [simple]\text{adj} \rightarrow [(\text{simple})\text{adj}]-\text{if}\text{y}]_v \rightarrow [\text{simplify}]_v \rightarrow [[\text{simplify}]_v-\text{cation}]_n \rightarrow [\text{simplification}]_n. Consequently, affixes could not select a base which is just characterized by containing an adjective at any depth of embedding, such as the hypothetical affix -\text{ix} which could attach to \text{simple}, \text{simplify}, \text{simplicity}, and \text{simplification}, but not to \text{drive}, \text{idea}, or \text{customize}. At the point where an affix would attach to \text{simplify} or \text{simplification}, the information that these words contain adjectives would simply be not available anymore. The central constraint on cyclic opacity in syntax is the Phase Impenetrability Condition (PIC, Chomsky 2001, Müller 2010), which states that syntactic heads above a Phase Head cannot target material below the ‘edge’ of the Phase head complement. The PIC has been argued to be of central importance not only for syntax, but also for morphophonology and semantics (see e.g., Marvin 2003) and allomorphy (Moskal 2015).

Where the PIC blocks processes across boundaries of an embedded cycle, the Strict Cycle Condition (SCC, Chomsky 1973, Mascaró 1976) blocks processes which are strictly inside of an embedded cycle, as in vowel elision in Emai (Casali 1997:512-13), where one of two consecutive vowels is deleted at the boundary of lexical words (/ko\text{ema/} ‘to plant yam’ \rightarrow [\text{kema}], but not inside of words (/\text{oao/} \rightarrow [\text{oao}]\text{[/]o}\text{a/} ‘the house’). This follows from the SCC under the assumption that words are cyclic domains. At the phrase level, /\text{oa/} is contained in an embedded phonological (word-level) cycle, and hence immune against the deletion rule. The SCC also plays a crucial role in cyclic models of syntax, blocking, e.g., movement in embedded clauses (Chomsky 1973, 1995). See Com\text{\textcircled{}}\text{Asp} for more discussion on the consequences of the SCC for the generative capacity of formal grammars.

A major tenet of syntactic research in the last decades was that PIC and SCC are not constraints, but follow from a derivational model where the relevant portions of phrase complements are made unaccessible by spellout, which would derive both restrictions. Similarly, Inkelas (1998) claims that Bracket Erasure falls out as a theorem of fine-grained cyclicity in sign-based phonology. The fundamental problem with this approach is that while there is substantial evidence for asymmetries corresponding to SCC and PIC, there is also a considerable body of exceptions (D’Alessandro & Scheer 2015, Newell 2018).

Bracket Erasure is thwarted by the phenomenon of potentiating (Orgun 1996). A central focus of project Syn\text{\textcircled{}}\text{Sem} are cases of semantic interpretation which go substantially beyond standard cyclic domains such as (extended) clauses. Against this background, a major question of the RU is whether constraints on cyclic opacity can be reinterpreted as viable constraints along the lines of McPherson & Heath (2016) and Sande & Jenks (2018), who argue that PIC effects should be captured by a viable faithfulness constraint for material from previous phases (Syn\text{\textcircled{}}\text{Mor}). A variant of this approach that we want to explore derives such effects more locally by gradient representations, where phases have the effect of adding phonological activation to processed phonological material, and faithfulness constraints react gradiently to iteratively computed phonological strength (Mor\text{\textcircled{}}\text{Phon}).

**D** **Heterogeneity of Cycles \approx Construction-specific Optimization**: A common question in cyclic and optimization approaches alike is how to handle cases where diferent domains in a language exhibit apparently different grammatical regularities (see also Com\text{\textcircled{}}\text{Asp} on the tight relation between Heterogeneity of Cycles and **B** Cyclic Timing). Cyclicity and optimization show a complementary potential in accounting for construction-specificity. Optimization allows for a principled account of domain-specific subregularities not available under cyclicity by itself. For example, variation in morpheme linearization has been a main incentive for optimization approaches in morphosyntax, specifically in the ordering
of clitics and affixes (Grimshaw 2001, Donohue 1998, Trommer 2001b, 2008, Zimmermann 2013a). (Optimizing) Repair rules investigated in detail in project Synphrase have built-in domain specifications as they allow linking different grammatical contexts with different possible repair strategies. Moreover, optimization allows for a principled account of cumulative construction-specificity (see Rosen 2016 and Trommer 2016b for representative analyses, and more discussion under project Morphrase).

On the other hand, cyclicity establishes principled locality domains for construction-specificity not provided by optimization on its own (Inkelas 1998, 2014) and captures a different type of construction specificity which seems to be systematically related to derivational depth by a mechanism that we may call cyclic aging: Representations change minimally by undergoing an earlier cycle $C_1$, which causes that they behave differently at a later cycle $C_2$ (see, e.g., Newell 2015 for a specific proposal). An important question we want to investigate in the RU is the added value from combining cycle-based and optimization-based approaches to construction-specificity. An implicit optimization component is already present in Newell’s age-based account where different languages allow different degrees of ‘invasiveness’ at later cycles. Another promising strategy is combining cycle-based Aging approaches and cumulative optimization of gradient representations (see Morphrase).
3 Bibliography


Bošković, Ž. 2014. Now I’m a phase, now I’m not a phase: On the variability of phases with extraction and ellipsis. LI 45(1). 27–89.


Müller, G. 2010. On Deriving CED Effects from the PIC. *LI* 41.1. 35–82.


