Computation also matters: a response to Pater (2018)*
(draft)

Adam Jardine, Rutgers University

June 6, 2018

Abstract

This article responds to Pater (2018) by arguing for a view of phonology that captures the computational properties of phonological processes. Jardine (2016)'s statement that tone is formally more complex than segmental phonology is not a claim, as Pater (2018) characterises it, but an empirical observation. This article thus outlines how phonological theories can incorporate such observations and integrate them with considerations of phonological substance. The conclusion is that, while computational characterisations are not necessarily alternatives to Optimality Theory, it is extremely difficult to capture the computational nature of phonological processes in Optimality Theory, due to the expressive power of global optimisation.

1 Introduction

Both substance and computation are central to a formal theory of phonology. *The Sound Pattern of English* (SPE; Chomsky and Halle, 1968) and Optimality Theory (OT; Prince and Smolensky, 2004) are both formal systems in that they specify

*Author contact: adam.jardine@rutgers.edu. This draft is indebted to invaluable discussion with Jeff Heinz and Bill Idsardi. All errors are my own.

computations over representations. What makes them phonological systems is that their representations—inputs, outputs, candidates, rules, constraints—refer to phonological "substance"; that is, representational primitives that are specific to phonology. However, because they are formal systems, they each make claims about the kinds of computations that are possible for the phonological module of the grammar. In SPE, this computation is handled by a rewrite rule formalism, and in OT the central computational mechanism is parallel optimisation through the generation and filtering of candidates.

The value of directly studying the computational properties of phonological patterns is thus to understand the kinds of computations that are performed by the phonological component of the grammar. Jardine (2016) identifies a computational difference between unbounded processes in segmental phonology and unbounded processes in tone, based on formal language-theoretic (FLT) characterisations of phonological processes as functions (Heinz and Lai, 2013; Chandlee, 2014). In other words, the space of possible computations for segmental phonology is strictly smaller than that of tonal phonology. Jardine (2016) argues that an OT theory cannot easily state this difference, citing examples of how optimisation does not adhere to the relevant computational properties.

In responding to Jardine 2016, Pater (2018) criticises these conclusions in two ways. First, he argues that the Jardine's FLT-based characterisations make incorrect predictions about the nature of tone. Second, he argues that FLT is "best understood not as an alternative to OT, but as a general tool for the formalisation and comparison of theories" (p. 152). In making this second point, he claims (p. 152) that FLT is unable to make statements about phonological "substance," where substance is defined broadly as both "restriction on combinations of formal primitives" and "the phonetic substance of phonology" (p. 155).

The goal of this paper is to rebut these arguments by clarifying the computational view of phonology. Importantly, what Pater (2018) cites as a "claim" (abstract; p. 151) that tone is formally more complex than segmental phonology is actually a *fact*. Segmental phonology inhabits an extremely restricted space of the range of possible functions (Heinz and Lai, 2013; Luo, 2017; Payne, 2017; Chandlee and Heinz, 2018), falling into what Heinz and Lai (2013) call the *weakly* *deterministic* class of functions. Jardine (2016) shows that processes outside of this space are well-attested in tone. If we accept this typological observation as a fact, then a complete theory of phonology ought to integrate some coherent statement of it.

A computational view of phonology encourages future phonological theories to build on this fact. Jardine (2016)'s findings are not a final theory of either tonal or segmental phonology, but instead point to important properties that should be maintained by theories of phonology. Thus, when Pater (2018) claims that Jardine's characterisations conflate high-tone spreading in Copperbelt Bemba (Bickmore and Kula, 2013, 2015) with truly unattested "sour grapes"-type spreading (Wilson, 2003, 2006), this misunderstands the nature of the argument. Whatever our theory of tone ultimately may be, the fact remains that tone commonly includes patterns that are computationally different from those in segmental phonology.

More detailed computational characterisations of segmental phonology are the subject of ongoing work (Chandlee et al., 2015; Chandlee and Heinz, 2018). These computational characterisations integrate the facts outlined by Jardine by positing futher restrictions on the weakly deterministic class. As detailed below, future work can also integrate substantive considerations into these facts through statements about the substructures and representations to which FLT-based grammars refer.

OT grammars also can state substantive considerations, but it is unclear whether they can incorporate the computational facts outlined by Jardine. The power of optimisation makes it difficult to reconcile OT grammars with the restrictive nature of phonological processes. Unbounded spreading is an excellent example—as Wilson (2003), McCarthy (2010), and Rose and Walker (2011) all point out, classic OT theories of spreading make incorrect typological predictions, to the extent that they have motivated significant deviations from the classic OT architecture. For this reason, while Pater (2018) claims to have shown how to capture Jardine (2016)'s generalisations in OT, it is not at all clear that these claims will withstand a full analysis of the factorial typology (a lá Alber et al. 2016). More importantly, as detailed below, classic OT's difficulties with spreading are directly related to the expressivity of the global evaluation available to OT. Reining in the power of optimisation thus requires a conspiracy of multiple, unrelated changes to the framework, making any coherent statement of Jardine (2016)'s observations extremely difficult.

In an important sense, Pater (2018) is correct when he states that FLT is not necesarily an alternative to OT. Computational characterisations of functions study the nature of the functions themselves, independent of how they may be generated intensionally. Thus, *a priori*, there is no reason that OT grammars are incompatible with FLT characterisations. In fact, any formal theory of phonology ought to be able to incorporate the facts made clear by these characterisations. Computational characterisations of phonology make clear typological and psycholinguistic predictions, and they help us to make progress on the learning problem (Heinz, 2018). Insofar as these are important goals for any phonological theory, FLT characterisations can serve as a guide for phonological theorizing. However, for the reasons outlined above, the predictions of OT theories of spreading do not square with the findings of FLT work on phonology.

Moving forward with phonological theory, it is crucial to take into account computational properties of phonological patterns. And it is by all means possible to reconcile these properties with considerations of phonological substance. A complete theory—which neither Jardine (2016) nor this paper claim to present—will need to do both. However, it appears that OT cannot, given that optimisation fails to capture the basic computational properties of phonological processes.

This paper is structured as follows. Section 2 briefly recapitulates the *weakly deterministic hypothesis*, the computational characterisation of segmental phonology around which Jardine (2016) builds his arguments, and Section 3 outlines how future work in theoretical phonology can build on this characterisation in formal terms. Section 4 outlines how to build on the weakly deterministic hypothesis by incorporating phonological substance, and compares and contrasts this to how substantive statements are made in OT. Section 5 then argues why the power of optimisation makes it difficult for OT to build on the weakly deterministic hypothesis. Section 6 discusses some exceptions to the weakly deterministic hypothesis, and how they are equally exceptional from the perspective of OT theories of spreading. Section 7 then concludes. The following discusses the issues in a brief and informal manner; interested readers are referred to the work cited below for more detailed explanations.

2 The weakly deterministic hypothesis

A phonological process can be studied as a function from an input string to an output string.[1] (For discussion of non-string representations, see §4.) We can then study phonological processes in terms of their structural properties as functions. A FLT characterisation of a function asks, "What kind of information is needed to compute this function?" These characterisations offer restrictive, precise, and testable hypotheses about the range of possible phonological processes.

Phonological processes are surprisingly limited in the information needed to compute them. In particular, the restrictive nature of segmental phonology is well-understood, even when taking unbounded processes into account. Segmental spreading has been characterized as myopic, meaning that it does not "look ahead" in the direction of spreading (Wilson, 2003, 2006; McCarthy, 2010). Similarly, it is well-established that long-distance consonantal harmony operates either right-to-left or propogates outward from a root (Hansson, 2001, 2010; Rose and Walker, 2004).

The *subsequential* class of functions fits the observed behaviour remarkably well. Subsequential functions are a strict subclass of the *regular* class of functions. The latter has been shown to be a reasonable upper bound for the complexity of phonological processes (Johnson, 1972; Kaplan and Kay, 1994; Frank and Satta, 1998). Briefly, a *left-subsequential* function is a regular function for which the output can be computed deterministically by processing the input string left to right. A *right-subsequential* function, likewise, can be computed deterministically from right to left. Taken together, these form the subsequential class of functions. The subsequential class has been studied in theoretical computer science

[1] This abstracts away from free variation, which is an important, yet distinct, problem. There are ways of characterizing processes with free variation that are extensions of the structural properties of functions discussed in this section (Mohri, 1997; Beros and de la Higuera, 2016).

(Schützenberger, 1977) and natural language processing (Mohri, 1997) for its naturalness and efficiency, and its specific structure gives it attractive learnability properties (Oncina et al., 1993).

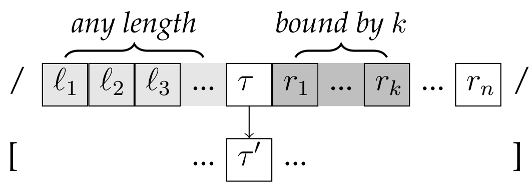
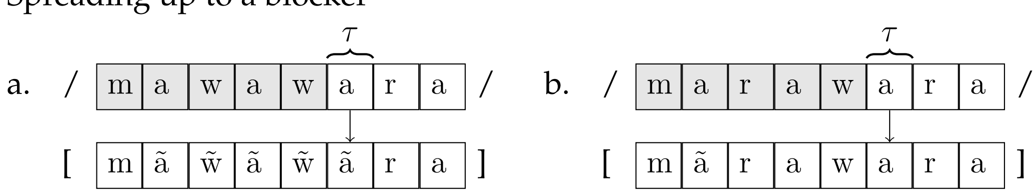


Figure 1: Computing a left-subsequential function. Here, τ is an arbitrary input target; τ' is its output. The sequence $\ell_1, \ell_2, \ell_3, ...$ in the input is the information available on the left side of τ available for computing τ', and $r_1, r_2, ..., r_k$ is the information available to the right.

Because the output of a subsequential function must be determined sequentially in a particular direction, then there must be some bound on how far ahead information used in determining that output can appear. This is illustrated in Fig. 1 for a left-subsequential function. For any target τ, its output can be determined by any amount of information to the left, but the amount of information to the right used in calculating its output must be bound by some constant k. The same is true for right-subsequential functions, except the bound is instead on the information to the left of the target (i.e., "ahead" of the scan, which is proceeding from right to left).

To briefly see how this applies to phonological processes, take an unbounded spreading process in which a nasal feature spreads left-to-right from a nasal consonant to all following vowels and glides, up to a non-nasal consonant (as in Johore Malay; Onn 1980).

(1) Spreading up to a blocker



When calculating the output of a target vowel, all information to the left of that vowel is important. If the target comes after a nasal consonant followed by a sequence of vowels and glides, the target surfaces as nasal (1a). If a liquid (and no other nasal) intervenes, then it surfaces as oral (1b). Note here that information to the right is irrelevant; thus, the bound k on lookahead to the right is 0, and so this spreading is left-subsequential.

A summary of bounds on processes and how that reflects on their status as functions is given in Table 1. Processes with a bound on both on the information to the left and to the right are *input strictly local* (Chandlee, 2014; Chandlee and Heinz, 2018). This is the case for "non-myopic" vowel harmony in Central Veneto (Walker, 2010), as both the left and right contexts of the process are bound by the position of stress. Processes with a bound only on the right context, like nasal spread in Johore Malay, are left-subsequential. Processes with a bound only on the left context, such as consonant assimilation in Inseño Chumash (Applegate, 1972; Hansson, 2010), are right-subsequential.

Position of bound	Type of Function
L & R contexts	input strictly-local
R context	left-subsequential
L context	right-subsequential
R context (parsing L→R), L context (parsing R→L)	weakly deterministic
None	beyond weakly det.

Table 1: Summary of types of function based on the boundedness of the contexts of a target (L=left, R=right). A dashed line indicates a region of complexity strictly more expressive than the region above it. The region highlighted in gray is hypothesized in (2) to be excluded in segmental phonology.

Thus, subsequentiality captures fundamental properties of the application of phonological processes. To accomodate bidirectional processes (such as root-control harmony), Heinz and Lai (2013) posit the *weakly deterministic* class. The weakly deterministic class extends the subsequential class to include functions consisting of the application of a single subsequential function both left-to-right and right-to-left. This class can be interpreted as a hypothesis about phonology, as stated in (2).

(2) *Weakly deterministic hypothesis* (Heinz and Lai, 2013; Jardine, 2016)

Segmental processes are weakly deterministic functions.

Typological studies have shown that local processes (Chandlee and Heinz, 2018), vowel harmony (Heinz and Lai, 2013), consonant harmony (Luo, 2017), and long-distance dissimilation (Payne, 2017) are all weakly deterministic, showing strong support for the hypothesis in (2). As Heinz and Lai (2013) point out, additional evidence for the weakly deterministic hypothesis is the absence of sour grapes, which they show to be not subsequential and conjecture to be not weakly deterministic. This is because sour grapes spreading requires unbounded lookahead both to the left and the right of any potential target (i.e., it lies in the bottom line of Table 1).

Jardine (2016), then, gives evidence that (2) must crucially apply only to segmental processes, given the common attestation of processes in tone that are not weakly deterministic. He cites two processes, unbounded tone plateauing (Kisseberth and Odden, 2003; Hyman, 2011) and unbounded spreading in Copperbelt Bemba (Bickmore and Kula, 2013, 2015) as examples of processes that require unbounded lookahead on *both* the left *and* the right of any given target. He then argues that phonological theory must be able to make direct reference to the weakly deterministic hypothesis in (2).

3 The weakly deterministic hypothesis and phonological explanation

Importantly, most functions are not weakly deterministic functions. That is, if we think about the universe of possible functions, even the universe of computable functions, the weakly deterministic class forms a tiny portion of this universe. Thus, (2) is an extremely *strong* hypothesis about the nature of segmental processes, because it predicts that all segmental processes fall in this small space.

This space has also been independently motivated in terms of processing and learnability.

However, even stronger hypotheses are possible. Not all weakly deterministic functions are attested segmental processes. For example, a process in which every third sibilant becomes [−anterior] would also be weakly deterministic, although it is not attested. Thus, the hypothesis in (2) is by no means the final word on phonology. However, further theorising should be a *restriction* on this hypothesis—that is, the predictions of future phonological theories should remain within the white area of Table 1.

One way to further restrict it is inherently computational—to examine subclasses of the weakly deterministic class. One example are the input- and output-strictly local functions, which are the subject of current research (Chandlee, 2014; Chandlee et al., 2015; Chandlee and Heinz, 2018). Another way is to add some sort of substantive constraints on the class; see §4. However, it is crucial to remember that all of these are *futher* restrictions on the weakly deterministic class—they do not change the fact that there is overwhelming empirical support for the weakly deterministic class as an upper bound for segmental phonology.

The same reasoning applies to tonal phonology. Clearly, the formal mechanisms available to tone need to account for non-weakly deterministic processes. However, they should still be at most regular, as even the non-weakly deterministic patterns Jardine cites are still regular functions. However, not all regular functions are attested tonal processes, and so there is room for even stronger characterisations of tone. Jardine (2016) explicitly leaves this as an open question, asking in his conclusion “What computational constraints are there on tone?” (p. 279).

Pater (2018) notes a difference between ‘true’ sour grapes and Copperbelt Bemba, and argues that the former is not attested in tone. If this is true, then our ultimate theory of phonology should exclude true sour grapes. The right theory may provide a substantive answer, as Pater suggests, or it may provide a computational one. Regardless, it does not change the fact that tone commonly includes non-weakly deterministic processes. Thus, Pater’s criticism of Jardine (2016) as not distinguishing between different non-weakly deterministic processes is misplaced.

4 Substantive explanation and the weakly deterministic hypothesis

Pater (2018) also assumes that any substance-based explanation cannot be reconciled with the weakly deterministic hypothesis: he states that FLT cannot provide "a means for the development of substantive theories of constraints" (p. 152). However, the weakly deterministic hypothesis in (2) is inherently a substantive one—it makes predictions about how phonological substance can interact in segmental versus other domains. In OT, statements about phonological substance are made through stipulations on what does and what does not appear in CON. Without a theory of CON or an analysis of its factorial typology, such stipulations are only descriptive restatements of the empirical facts in terms of the framework, not a predictive theory of phonology.

The weakly deterministic hypothesis *predicts* how phonological substance will interact. There is no real difference between tone and segmental phonology in the substance Pater refers to—in terms of positional licensing, both tone and vowel features have been shown to refer to prominent positions. What the FLT analysis of their difference *explains* is that vowel features interact with that substance in a different way than tone.

Pater (2018) claims to have explained this in OT, but his paper gives neither a "theory of constraints" nor a real analysis of the typological predictions of the particular constraint set he advocates for. A "theory of constraints" requires an explicit constraint definition language (Eisner, 1997; de Lacy, 2011) that governs the constraint set. Without a principled theory of CON, it is merely a stipulation that the licensing constraint φ_{fin}-H motivating non-myopic spreading in Copperbelt Bemba appears in CON but no such constraint using vowel features appears (a fact that Pater admits; p. 154).

Furthermore, a particular choice for CON is only successful as a theory insofar as its factorial typology has been sufficiently analysed (Alber et al., 2016).

This issue is particularly acute for constraints designed to capture unbounded spreading, which, as to be detailed in the following section, are well-known for their problematic typological predictions. Thus, counter to his claim, Pater (2018) has not shown that a purely substantive theory predicts the typological facts of unbounded spreading. Instead, he has simply restated, in terms of OT, the fact that CB appears in tone but not in segmental phonology.

Additionally, Pater 2018 claims that FLT "provides no obvious way of stating the kinds of substantive restrictions" that are useful in capturing phonological generalisations (p. 156), but this is also not the case. Substance can be incorporated into FLT grammars through restrictions on the constraints grammars can choose from, and through the representations over which grammars operate. In other words, one can make substantive stipulations in FLT grammars much in the same way they are made in OT, or in any other theory of phonology.

This is most obviously true for phonotactic grammars based in FLT, which capture patterns through constraints that specifiy illicit substructures (Heinz, 2010; McMullin and Hansson, 2016; Jardine and Heinz, 2016), just as markedness constraints in OT (de Lacy, 2011). The only difference is that FLT constraints are interpreted as inviolable. Thus, just as substantive statements can be made in OT by stipulating the content of CON, we can similarly make substantive statements in FLT phonotactic grammars through stipulations on what constraints are available to grammars. This is thus also true for processes, as we can study classes of functions in terms of the range of output patterns they generate (see, e.g., Chandlee and Heinz, 2018). It is also possible to study the range of functions that perform repairs on particular marked structures (Chandlee et al., 2015), so here we can also make substantive statements by specifying which structures are marked and need to be repaired.

Furthermore, as with other kinds of phonological grammars, substance can also be incorporated into phonological theories based in FLT through representation. This is implicit in the alphabets of symbols chosen in FLT-based work—Heinz (2009) uses strings of syllables to characterize stress, Heinz and Lai (2013) assume a vowel tier, and Jardine (2016) uses strings of moras. FLT-based studies of non-string phonological representations are the subject of ongoing research;

see Jardine (2017) for tonal autosegmental representations and Strother-Garcia (2017) for syllable representations. In particular, the reader is referred to Jardine (2018), who argues that formal grammars defined over autosegmental representations more naturally capture the typology of tone well-formedness patterns than established formal grammars over strings.

Finally, Pater (2018) states that "[i]t is worth noting that the formal complexity statement leaves completely unaddressed why tonal systems should be more complex." (p. 156). This is not entirely true. In fact, Jardine (2016) makes an appeal to the same prosodic substance that Pater does: "It could be that tone has access to such resources because prosodic information more commonly interacts with syntax ... and thus requires more powerful computation" (p. 276). Admittedly, a full understanding of how this comes to be requires a more articulated theory of processing and computation at the syntax-phonology interface. While such a theory may be far off, it will be more meaningful than a simple statement of whether or not a constraint appears in CON.

5 Formal explanation and Optimality Theory

Thus, substantive considerations can be incorportated into computational theories of phonology. However, computational properties are not easily incorporated into Optimality Theory, due to the computational power of optimisation that is evaluated globally.

Pater (2018) states that "most OT theories of spreading do not produce sour grapes" (p. 154), however the pathological predictions of parallel OT with regards to unbounded spreading are well-documented (Wilson, 2003, 2006; McCarthy, 2010; Rose and Walker, 2011). It is notable that, to capture the myopia generalisation, both Wilson (2003, 2006) and McCarthy (2010) appeal to significant changes to how OT evaluates candidates. Thus, Pater's claim that he has demonstrated an OT system that "can generate the desired typology" (p. 155) is not sufficiently supported.

The reason for these issues is clear when we examine the expressive power of optimisation. Even with simple constraints, OT grammars can compute nonsubsequential, fully regular functions such as sour grapes (Heinz and Lai, 2013) as well as non-regular functions such as majority rules (Gerdemann and Hulden, 2012; Heinz and Lai, 2013) and alphabetical sorting (Lamont, 2018).

To give an example, sorting can be produced by ALIGN-based spreading constraints. Take rightward spreading of a [nasal] feature, as discussed in §2. As pointed out by McCarthy (2010), ranking an ALIGN-R([nasal],word) constraint over LINEARITY predicts metathesis of a blocker and undergoer in order to more optimally align the [nasal] feature. In fact, the problem is worse than McCarthy illustrates: given a string of undergoers and blockers, *all* blockers are sent to the right edge of the word in a 'mass-metathesis' so that the winning candidate sorts the form into a nasal portion followed by a non-nasal portion.

(3)

	/mawararara/	*NASLIQ	ALIGN-R([nas])	LIN	ID([nas])
☞	mãw̃ããããrrr	0	3	6	6
	mãw̃ããrrara	0	5	1	4
	mãw̃ãrarara	0	6	0	3
	mãw̃ãr̃ãr̃ãr̃ã	3	0	0	9

The winning candidate in (3) is [mãw̃ããããrrr], in which LINEARITY has been massively violated so that all of the undergoers are ordered before blockers. This is not even a regular function, because the amount of memory required to compute the output grows with the number of blockers in the input that must be deposited at the end of the word. Thus, while there are OT theories of spreading that do not produce sour grapes, they can still miss the generalisation that segmental spreading is weakly deterministic (and even that it is regular).

Issues like sour grapes and mass-metathesis arise because optimisation allows direct comparison of candidates with local changes to candidates with non-local changes. This 'global' evaluation of optimisation is, as these examples show, computationally very powerful. It may be possible to rein in the power of OT grammars by carefully choosing the right constraints. We could reject AGREE, as suggested by Pater (2018). We would also have to reinterpret LINEARITY as an inviolable part of GEN, as the violability of LINEARITY is partially the culprit for the pathology in (3). However, all such modifications 'conspire' to capture a generalisation that is stated directly by the weakly deterministic hypothesis.

Furthermore, it does not appear to be the case that OT's problems with unbounded spreading can be solved by simply choosing the right constraint set. In their review of harmony systems and their analyses, Rose and Walker (2011) tellingly note: "Proposals like those made by Wilson [to solve overgeneration problems with spreading constraints –AJ] ... involve substantial departures from traditional constraint architecture in OT" (p. 268). The reason why Wilson (2003) proposes targeted constraints, which only assess a limited set of candidates, is because he finds that previous analyses of harmony in classical OT fail to capture the basic nature of segmental spreading. Weak determinism, in contrast, captures it directly, and provides a meta-theoretical bound for future theories to adhere to.

6 Exceptions to the weakly deterministic hypothesis

It should be noted that non-weakly deterministic segmental processes, although extremely rare, are attested. One, discussed in Jardine (2016), is an apparent unbounded plateauing vowel harmony process in Yaka (Hyman, 1998). Another, which has come to light since Jardine (2016) was published, is ATR harmony in Tutrugbu (McCollum and Essegbey, 2018), in which unbounded harmonizing of low vowels to a following [+ATR] vowel is blocked by a preceding [+high] initial prefix.

Such patterns are of interest because they challenge a categorical interpretation of the weakly deterministic hypothesis. Indeed, McCollum et al. (2017) interpret the facts of Tutrugbu as evidence for the weakly deterministic bound as a bias, and not a categorical constraint (see also Avcu, 2018). However, it is also the case that neither are evidence for Pater (2018)'s OT-based prosodic licensing explanation of non-myopic processes. Hyman's analysis for Yaka invokes a sour-grapes style analysis, and McCollum and Essegbey invoke a combination of spreading and correspondence constraints. Thus, both processes are just as unexpected for OT-based theories of spreading. The upshot is that both cases warrant further investigation, regardless of one's theoretical perspective.

7 Conclusion

Phonological processes share non-arbitrary structural properties. A theory of phonology should thus be able to state these properties directly. Segmental processes are, almost without exception, weakly deterministic. This generalisation escapes explanation in OT because, as illustrated in §5, global optimisation makes it difficult, if not impossible, to make any coherent statement about the structural properties of phonological processes in OT.

FLT characterisations of phonology nontrivially capture generalisations that are important to any theory of phonology. At the least, then, they can serve as constraints on the kinds of computations that phonological theories should allow. Thus, when we look to include substantive considerations, we should do in a way that is in harmony with computational characterisations. Given the challenges of reining in optimisation, it is not clear that this is possible in OT.

References

Alber, Birgit, DelBusso, Natalie, and Prince, Alan (2016). From intensional properties to universal support. *Language: Phonological Analysis*, 92. e88–116.

Applegate, Richard B. (1972). *Ineseño Chumash grammar*. Ph.D. dissertation, University of California, Berkeley.

Avcu, Ennis (2018). Experimental investigation of subregular hypothesis. In *WCCFL 35*, 77–86.

Beros, Achilles and de la Higuera, Colin (2016). A canonical semi-deterministic transducer. *Fundamenta Informaticae*, 431–459.

Bickmore, Lee S. and Kula, Nancy C. (2013). Ternary spreading and the OCP in Copperbelt Bemba. *Studies in African Linguistics*, 42.

Bickmore, Lee S. and Kula, Nancy C. (2015). Phrasal phonology in Copperbelt Bemba. *Phonology*, 32. 147–176.

Chandlee, Jane (2014). *Strictly Local Phonological Processes*. Ph.D. dissertation, University of Delaware.

Chandlee, Jane and Heinz, Jeffrey (2018). Strictly locality and phonological maps. *LI*, 49. 23–60.

Chandlee, Jane, Jardine, Adam, and Heinz, Jeffrey (2015). Learning repairs for marked structures. In *Proceedings of the 2015 Annual Meeting on Phonology*. LSA.

Chomsky, Noam and Halle, Morris (1968). *The Sound Pattern of English*. Harper & Row.

de Lacy, Paul (2011). Markedness and faithfulness constraints. In Marc Van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice (eds.), *The Blackwell Companion to Phonology*. Blackwell.

Eisner, Jason (1997). What constraints should OT allow? Talk handout, Linguistic Society of America, Chicago. ROA#204-0797. Available at http://roa.rutgers.edu/.

Frank, Robert and Satta, Giorgio (1998). Optimality theory and the generative capacity complexity of constraint violability. *Computational linguistics*, 24. 277–299.

Gerdemann, Dale and Hulden, Mans (2012). Practical finite state optimality theory. In *Proceedings of the 10th International Workshop on FSMNLP*, 10–19. ACL.

Hansson, Gunnar Ólafur (2001). *Theoretical and Typological Issues in Consonant Harmony*. Ph.D. dissertation, University of California, Berkeley.

Hansson, Gunnar Ólafur (2010). *Consonant Harmony: Long-distance Interaction in Phonology*. Berkeley: University of California Press.

Heinz, Jeffrey (2009). On the role of locality in learning stress patterns. *Phonology*, 26. 303–351.

Heinz, Jeffrey (2010). Learning long-distance phonotactics. *LI*, 41. 623–661.

Heinz, Jeffrey (2018). The computational nature of phonological generalizations. In Larry Hyman and Frans Plank (eds.), *Phonological Typology*, Phonetics and Phonology, chapter 5, 126–195. De Gruyter Mouton.

Heinz, Jeffrey and Lai, Regine (2013). Vowel harmony and subsequentiality. In Andras Kornai and Marco Kuhlmann (eds.), *Proceedings of the 13th Meeting on Mathematics of Language*. Sofia, Bulgaria.

Hyman, Larry (1998). Positional prominence and the 'prosodic trough' in Yaka. *Phonology*, 15. 14–75.

Hyman, Larry (2011). Tone: Is it different? In John A. Goldsmith, Jason Riggle, and Alan C. L. Yu (eds.), *The Blackwell Handbook of Phonological Theory*, 197–238. Wiley-Blackwell.

Jardine, Adam (2016). Computationally, tone is different. *Phonology*, 33. 247–283.

Jardine, Adam (2017). The local nature of tone-association patterns. *Phonology*, 34. 385–405.

Jardine, Adam (2018). The expressivity of autosegmental grammars. *Journal of Logic, Language, and Information*, (Online first). 1–46. ISSN 1572-9583. doi:10.1007/s10849-018-9270-x.

Jardine, Adam and Heinz, Jeffrey (2016). Learning tier-based strictly 2-local languages. *Transactions of the Association for Computational Linguistics*, 4. 87–98. ISSN 2307-387X.

Johnson, C. Douglas (1972). *Formal aspects of phonological description*. Mouton.

Kaplan, Ronald and Kay, Martin (1994). Regular models of phonological rule systems. *Computational Linguistics*, 20. 331–78.

Kisseberth, Charles and Odden, David (2003). Tone. In Derek Nurse and Gérard Philippson (eds.), *The Bantu Languages*. New York: Routledge.

Lamont, Andrew (2018). Precedence is pathological: The problem of alphabetical sorting. Poster, WCCFL 36.

Luo, Huan (2017). Long-distance consonant agreement and subsequentiality. *Glossa*, 2:52.

McCarthy, John (2010). Autosegmental spreading in Optimality Theory. In John A. Goldsmith, Elizabeth Hume, and W. Leo Wezels (eds.), *Tones and Features: Phonetic and Phonological Perspectives*, 195–222. De Gruyer Mouton.

McCollum, Adam, c, Eric Bakovi; Mai, Anna, and Meinhardt, Eric (2017). Conditional blocking in Tutrugbu requires non-determinism: implications for the subregular hypothesis. In *Paper presented at NELS 48*.

McCollum, Adam G. and Essegbey, James (2018). Unbounded harmony is not always myopic: Evidence from Tutrugbu. In Wm. G. Bennett, Lindsay Hracs, and Dennis Ryan Storoshenko (eds.), *WCCFL 35*, 251–258.

McMullin, Kevin and Hansson, Gunnar Ólafur (2016). Long-distance phonotactics as Tier-Based Strictly 2-Local languages. In *Proceedings of AMP 2015*.

Mohri, Mehryar (1997). Finite-state transducers in language and speech processing. *Computational Linguistics*, 23:2. 269–311.

Oncina, José, García, Pedro, and Vidal, Enrique (1993). Learning subsequential transducers for pattern recognition tasks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15. 448–458.

Onn, Farid M. (1980). *Aspects of Malay Phonology and Morphology: A Generative Approach*. Kuala Lampur: Universiti Kebangsaan Malaysia.

Pater, Joe (2018). Substance matters: a reply to Jardine (2016). *Phonology*, 35:1. 151–156.

Payne, Amanda (2017). All dissimilation is computationally subsequential. *Language: Phonological Analysis*, 93:4. e353–371.

Prince, Alan and Smolensky, Paul (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell Publishing.

Rose, Sharon and Walker, Rachel (2004). A typology of consonant agreement as correspondence. *Language*, 80. 475–531.

Rose, Sharon and Walker, Rachel (2011). Harmony systems. In John Goldsmith, Jason Riggle, and Alan Yu (eds.), *The Handbook of Phonological Theory*, 240–290. Wiley-Blackwell.

Schützenberger, Marcel Paul (1977). Sur une variante des fonctions séquentielles. *Theoretical Computer Science*, 4. 47–57.

Strother-Garcia, Kristina (2017). Imdlawn Tashlhiyt Berber syllabification is quantifier-free. In *Proceedings of the first annual meeting of the Society for Computation in Linguistics*, volume 1, 145–153.

Walker, Rachel (2010). Nonmyopic harmony and the nature of derivations. *LI*, 41. 169–179.

Wilson, Colin (2003). Analyzing unbounded spreading with constraints: marks, targets, and derivations. Unpublished ms.

Wilson, Colin (2006). Unbounded spreading is myopic. Paper presented at the workshop Current Perspectives on Phonology.