

Formal Parameters of Phonology^{*}

From Government Phonology to SPE

Thomas Graf

Department of Linguistics
University of California, Los Angeles
tgraf@ucla.edu
<http://tgraf.bol.ucla.edu>

Abstract. Inspired by the model-theoretic approach to phonology deployed by Kracht [25] and Potts and Pullum [32], I develop an extendable modal logic for the investigation of phonological theories operating on (richly annotated) string structures. In contrast to previous research in this vein [17, 31, 37], I ultimately strive to study the entire class of such theories rather than merely one particular incarnation thereof. To this end, I first provide a formalization of classic Government Phonology in a restricted variant of temporal logic, whose generative capacity is then subsequently increased by the addition of further operators, thereby pushing it up the subregular hierarchy until one reaches the level of the regular stringsets. I identify several other axes along which Government Phonology might be generalized, moving us towards a parametric metatheory of phonology.

Like any other subfield of linguistics, phonology is home to a multitude of competing theories that differ vastly in their conceptual and technical assumptions. Contentious issues are, among others, the relation between phonology and phonetics (and if it is an interesting research question to begin with), if features are privative, binary or attribute valued, if phonological structures are strings, trees or complex matrices, if features can move from one position to another (i.e. if they are *autosegments*), and what role optimality requirements play in determining well-formedness. Meticulous empirical comparisons carried out by linguists have so far failed to yield conclusive results; it seems that for every phenomenon that lends support to a certain set of assumptions, there is another one that refutes it.

The lack of a theoretical consensus should not be taken to indicate that the way phonologists go about their research is flawed. Unless one subscribes to the view that scientific theories can faithfully reflect reality rather than merely approximate it, it is to be expected that one theory may fail where another one succeeds, and *vice versa*. A similar situation arises in physics, where depending

^{*} This paper has benefited tremendously from the comments and suggestions of Bruce Hayes, Ed Keenan, Marcus Kracht, Ed Stabler, Kie Zuraw, the members of the UCLA phonology seminar (winter quarter 2009), and two anonymous reviewers.

on the circumstances light exhibits particle-like or wave-like properties. But faced with this apparent indeterminacy of theory choice, it is only natural for us to ask if there is a principled way to identify interchangeable theories, i.e. proposals which may seem to have little in common yet are underlyingly the same. This requires developing a metatheory of phonology that uses a finite set of parameters to conclusively determine the equivalence class which a given phonological theory belongs to. This paper is intended to lay the basis for such a metatheory, building on techniques and insights from model-theoretic syntax [24, 35, 36]: I develop a modal logic for the formalization of a particular theory, Government phonology (GP), and then use this modal logic and its connections to neighboring areas, foremost formal language theory, to explore natural extensions and their relation to other approaches in phonology.

I feel obliged to point out in advance that I have my doubts concerning the feasibility of a formal theory of phonology that is adequate and insightful on both a linguistic and a mathematical level. But this is a problem all too familiar to mathematical linguists: any mathematically natural class of formal languages allows for constructions that never arise in natural language. For example, assignment of primary word stress is sometimes sensitive to whether a syllable is an odd or an even number of syllables away from the edge of a word (see [10] and my remarks in Sec. 2). Now in order to distinguish between odd and even, phonology has to be capable of counting modulo 2. On the other hand, phenomena that involve counting modulo 3, 4 or 21 — which from a mathematical perspective are just as simple as counting modulo 2 — are unheard of. Thus, the problem of mathematical methods in the realm of language is that their grip tends to be too loose, and the more we try to tighten it, the more difficult it becomes to prove interesting results.

Undeniably, though, a loose grip is better than no grip at all. I am confident that in attempting to construct the kind of metatheory of phonology I envision, irrespective of any shortcomings it might have, we will gain crucial insights into the core claims about language that are embodied by different phonological assumptions (e.g. computational complexity and memory usage) and how one may translate those claims from one theory into another. Moreover, the explicit logical formalization of linguistic theories makes it possible to investigate various problems in an algorithmic way using techniques from proof theory and model checking. These results are relevant to linguists and computer scientists alike. Linguists get a better understanding of how their claims relate to the psychological reality of language, how the different modules of a given theory interact to yield generalizations, and how they increase the expressivity of a theory (see [32] for such results on optimality theory). To a limited degree, linguists also get the freedom to switch to different theories for specific phenomena without jeopardizing the validity of their framework of choice. Computer scientists, on the other hand, will find that the model-theoretic perspective on phonology eases the computational implementation of linguistic proposals and allows them to gauge their runtime-behavior in advance. Furthermore, they may use the connection between finite model theory and formal language theory to increase the

efficiency of their programs by picking the weakest phonological theory that is expressive enough for the task at hand.

This paper is divided into two parts as follows. First, I introduce GP as an example of a weak theory of phonology and show how it can be axiomatized as a theory of richly annotated string structures using modal logic. In the second part, I analyze several parameters that distinguish GP from other proposals and might have an effect on generative capacity. In particular, I discuss how increasing the power of GP’s spreading operation moves us along the subregular hierarchy and why the specifics of the feature system have no effect on expressivity in general. I close with a short discussion of two important areas of future research, the impact of the syllable template on generative capacity and the relation between derivational and representational theories.

The reader is expected to have some basic familiarity with phonology, formal language theory, non-classical logics and model-theoretic syntax. There is an abundance of introductory material for the former three, while the latter is cogently summarized in [34] and [35].

1 A Weak Theory of Phonology — Government Phonology

1.1 Informal Overview

Due to space restrictions, I offer but a sketch of the main ideas of Government Phonology (GP). More readily accessible expositions may be found in the User’s Guide to Government Phonology [20] and related work of mine [10, 11]. To compensate for the terseness, the reader may want to check the explanation against the examples in Fig. 1 on the following page. Before we go *in medias res*, though, a note on my sources is in order. Just like Government-and-Binding theory [4], GP has changed a lot since its inception and practitioners hardly ever fully specify the details of the version of GP they use. However, there seems to be a consensus that a GP-variant is considered canonical if it incorporates the following modules: government, the syllable template, coda licensing and the ECP from [21], magic licensing from [19], and licensing constraints and the revised theory of elements from [20]. My strategy will be to follow the definitions in [20] as closely as possible and fill in any gaps using the literature just cited.

In GP, the carrier of all phonological structure is the *skeleton*, a finite, linearly ordered sequence of nodes (depicted by little crosses in Fig. 1) to which phonological expressions (PEs) can be attached in order to form the *melody* of the structure. A PE is built from a set E of privative features called *elements*, yielding a pair $\langle O, H \rangle$, where $O \subseteq E$ is a set of *operators*, $H \in E \cup \{\emptyset\}$ the *head*, and $H \notin O$. It is an open empirical question how many features are needed for an adequate account of phonological behavior [13, 14] — recent incarnations usually set $E := \{A, I, U, H, L, ?\}$, but for our axiomatization the only requirement is for E to be finite. Some examples of PEs are $[s] = \langle \{A, H\}, \emptyset \rangle$, $[n] = \langle \{L, ?\}, A \rangle$, $[i] = \langle \emptyset, \emptyset \rangle$, $[ɪ] = \langle \{I\}, \emptyset \rangle$, $[i] = \langle \emptyset, I \rangle$, and $[j] = \langle \emptyset, I \rangle$. The set of licit PEs is

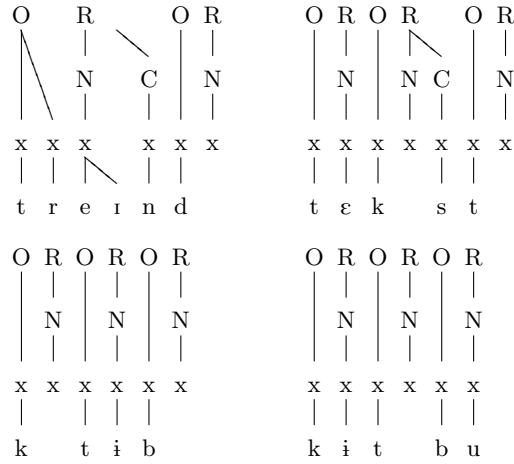


Fig. 1. Some phonological structures in GP (with IPA notation)

further restricted by language-specific *licensing constraints*, i.e. restrictions on the co-occurrence of features and their position in the PE. Common licensing constraints are for A to occupy only head positions, ruling out [s] in the list above, and for I and U not to occur in the same PE, ruling out the typologically uncommon $[y] = \langle \{U\}, I \rangle$ and $[ɣ] = \langle \{I\}, U \rangle$, among others.

As witnessed by $[i] = \langle \emptyset, I \rangle$ and $[j] = \langle \emptyset, I \rangle$, every PE is inherently underspecified; whether it is realized as a consonant or a vowel depends on its position in the structure, which is annotated with constituency information. An expression is realized as a vowel if it is associated to a skeleton node contained by a *nucleus* (N), but as a consonant if the node is contained by an *onset* (O) or a *coda* (C). Every N constitutes a *rhyme* (R), with C an optional subconstituent of R. All O, N and R may *branch*, that is be associated to up to two skeleton nodes, but a branching R must not contain a branching N. Furthermore, word initial O can be *floated*, i.e. be associated to no node at all. The number of PEs per node is limited to one, with the exception of unary branching N, where the limit is two (to model light diphthongs).

All phonological structures are obtained from concatenating $\langle O, R \rangle$ pairs according to constraints imposed by two government relations. *Constituent government* restricts the distribution of elements within a constituent, requiring that the leftmost PE licenses all other constituent-internal PEs. *Transconstituent government* enforces dependencies between the constituents themselves. In particular, every branching O has to be licensed by the N immediately following it, and every C has to be licensed by the PE contained in the immediately following O. Even though the precise licensing conditions are not fully worked out for either government relation, the general hypothesis is that PE_i licenses PE_j iff PE_i is leftmost in its constituent and contained by N, or leftmost in its constituent and composed from at most as many elements as PE_j and licenses no $PE_k \neq PE_j$

(hence any C has to be followed by a non-branching O, but a branching O might be followed by a branching N or R).

GP also features *empty categories*: a segment does not have to be associated to a PE. Inside a unary branching O, an unassociated node will always be mapped to the empty string. Inside N, on the other hand, it is either mapped to the empty string or the language-specific realization of the PE $\langle\{\emptyset\}, \emptyset\rangle$. This is determined by the phonological ECP, which allows only *p-licensed* N to be mapped to the empty string. N is p-licensed if it is followed by a coda containing a sibilant (*magic licensing*), or in certain languages if it is the rightmost segment of the string (*final empty nucleus*, abbreviated FEN), or if it is *properly governed* [18]. N is properly governed if the first N following it is not p-licensed and no government relations hold between or within any Cs or Os in-between the two Ns. Note that segments inside C or a branching O always have to be associated to a PE.

Finally, GP allows elements to *spread*, just as in fully autosegmental theories [9]. All elements, though, are assumed to share a single tier, and association lines are allowed to cross. The properties of spreading have not been explicitly spelled out in the literature, but it is safe to assume that it can proceed in either direction and might be optional or obligatory, depending on the element, its position in the string and the language in question. While there seem to be restrictions on the set of viable targets given a specific source, the only canonical one is a ban against spreading within a branching O.

1.2 Formalization in Modal Logic

For my formalization, I use a very weak modal logic that can be thought of as the result of removing the “sometime in the future” and “sometime in the past” modalities from restricted temporal logic [6, 7]. Naturally, the tree model property of modal logic implies that the logic is too weak to define the intended class of models, so we are indeed dealing with a formal description rather than a proper axiomatization.

Let E be some non-empty finite set of *basic elements* different from the neutral element v , which represents the empty set of GP’s feature calculus. We define the set of *elements* $\mathcal{E} := (E \times \{1, 2\} \times \{\text{head}, \text{operator}\} \times \{\text{local}, \text{spread}\}) \cup (\{v\} \times \{1, 2\} \times \{\text{head}, \text{operator}\} \times \{\text{local}\})$. The intended role of the *head/operator* and *local/spread* parameter is to distinguish elements according to their position in the PE and whether they arose from a spreading operation, respectively. The second projection is of very limited use and required only by GP’s rendition of light diphthongs as two PEs associated to one node in the structure. The set of melodic features $\mathcal{M} := \mathcal{E} \cup \{\mu, \text{fake}, \checkmark\}$ will be our set of propositional variables. The intention is for μ (mnemonic for *mute*) and \checkmark to mark unpronounced and licensed segments, respectively, while *fake* denotes an unassociated onset. For the sake of increased readability, the set of propositional variables is “sorted” such that $x \in \mathcal{M}$ is represented by m , $m \in \mathcal{E}$ by e , heads by h , and operators by o . The variable e_n is taken to stand for any element such that $\pi_2(e) = n$, where

$\pi_i(x)$ returns the i^{th} projection of x . In rare occasions, I will write \underline{e} and \bar{e} for a specific element e in head and operator position, respectively.

Furthermore, there are three nullary modalities¹, N , O , C , the set of which is designated by \mathcal{S} , read *skeleton*. In addition, we introduce two unary diamond operators \triangleleft and \triangleright , whose duals are denoted by \blacktriangleleft and \blacktriangleright . The set of well-formed formulas is built up in the usual way from \mathcal{M} , \mathcal{S} , \triangleleft , \triangleright , \rightarrow and \perp .

Our intended models $\mathfrak{M} := \langle \mathfrak{F}, V \rangle$ are built over bidirectional frames $\mathfrak{F} := \langle D, R_i, R_{\triangleleft} \rangle_{i \in \mathcal{S}}$, where D is an initial subset of \mathbb{N} , $R_i \subseteq D$ for each $i \in \mathcal{S}$, and R_{\triangleleft} is the successor function over \mathbb{N} . The valuation function $V : \mathcal{M} \rightarrow \wp(D)$ maps propositional variables to subsets of D . The definition of satisfaction is standard, though it should be noted that our models are “numbered from right to left”. That is to say, $0 \in D$ marks the right edge of a structure and $n + 1$ is to the left of n . This is due to GP’s transconstituent government being computed from right to left.

$$\begin{aligned}
\mathfrak{M}, w \models \perp & \quad \text{never} \\
\mathfrak{M}, w \models p & \quad \text{iff } w \in V(p) \\
\mathfrak{M}, w \models \neg\phi & \quad \text{iff } \mathfrak{M}, w \not\models \phi \\
\mathfrak{M}, w \models \phi \wedge \psi & \quad \text{iff } \mathfrak{M}, w \models \phi \text{ and } \mathfrak{M}, w \models \psi \\
\mathfrak{M}, w \models N & \quad \text{iff } w \in R_N \\
\mathfrak{M}, w \models O & \quad \text{iff } w \in R_O \\
\mathfrak{M}, w \models C & \quad \text{iff } w \in R_C \\
\mathfrak{M}, w \models \triangleleft \phi & \quad \text{iff } \mathfrak{M}, w + 1 \models \phi \\
\mathfrak{M}, w \models \triangleright \phi & \quad \text{iff } \mathfrak{M}, w - 1 \models \phi
\end{aligned}$$

With the logic fully defined, we can turn to the axioms for GP. The formalization of the skeleton is straightforward if one models binary branching constituents as two adjacent unary branching ones and views rhymes as mere notational devices. Recall that Ns containing light diphthongs are implemented as a single N with both e_1 and e_2 elements associated to it.

S1	$\bigwedge_{i \in \mathcal{S}} (i \leftrightarrow \bigwedge_{i \neq j \in \mathcal{S}} \neg j)$	Unique constituency
S2	$(\blacktriangleleft \perp \rightarrow O) \wedge (\blacktriangleright \perp \rightarrow N)$	Word edges
S3	$R \leftrightarrow (N \vee C)$	Definition of rhyme
S4	$N \rightarrow \triangleleft O \vee \triangleleft N$	Nucleus placement
S5	$O \rightarrow \neg \triangleleft O \vee \neg \triangleright O$	Binary branching onsets
S6	$R \rightarrow \neg \triangleleft R \vee \neg \triangleright R$	Binary branching rhymes
S7	$C \rightarrow \triangleleft N \wedge \triangleright O$	Coda placement

¹ I follow the terminology of [1] here. Nullary modalities correspond to unary relations and can hence be thought of as propositional constants. As far as I can see, nothing hinges on whether we treat constituent labels as nullary modalities, propositional constants, or propositional variables; my motivation in separating them from phonological features stems solely from the parallel distinction between melody and constituency in GP.

GP’s feature calculus is also easy to capture. A propositional formula ϕ over a set of variables x_1, \dots, x_k is called *exhaustive* iff $\phi := \bigwedge_{1 \leq i \leq k} \psi_i$, where for every i , ψ_i is either x_i or $\neg x_i$. A PE ϕ is an exhaustive propositional formula over \mathcal{E} such that $\phi \cup \{\mathbf{F1}, \mathbf{F2}, \mathbf{F3}, \mathbf{F4}, \bigvee h, \bigvee o\}$ is consistent.

F1	$\bigwedge (h_n \rightarrow \bigwedge_{h_n \neq h'_n} \neg h'_n)$	Exactly one head
F2	$\neg v \rightarrow \bigwedge (h_n \rightarrow \bigwedge_{\pi_1(h)=\pi_1(o)} \neg o_n)$	No basic element (except v) twice
F3	$\bar{v} \rightarrow \bigwedge_{o \neq \bar{v}} \neg o$	\bar{v} excludes other operators
F4	$\bigwedge (e_2 \rightarrow \bigvee h_1 \wedge \bigvee o_1)$	Pseudo branching implies first branch

Let PH be the least set containing all PEs (noting that a PE is now a particular kind of propositional formula), and let $lic : PH \rightarrow \wp(PH)$ map every PE to its set of melodic licensors. Furthermore, $S \subseteq PH$ designates the set of PEs occurring in the codas of magic licensing configurations (the letter S is mnemonic for “sibilants”). The following five axioms, then, sufficiently restrict the melody.

M1	$\bigwedge_{i \in S} \left(i \rightarrow \bigvee_{\phi \in PH} \phi \vee \mu \vee \mathbf{fake} \right)$	Universal annotation
M2	$((O \vee \triangleleft N \vee \triangleright N) \rightarrow \bigwedge \neg e_2)$	No pseudo branching for O, C & branching N
M3	$O \wedge \triangleleft O \rightarrow \bigwedge_{\phi \in PH} (\phi \rightarrow \bigvee_{\psi \in lic(\phi)} \triangleleft \psi)$	Licensing within branching onsets
M4	$C \wedge \bigwedge_{i \in S} \neg i \rightarrow \triangleleft \neg \mu \wedge \bigwedge_{\phi \in PH} (\phi \rightarrow \bigvee_{\psi \in lic(\phi)} \triangleright \psi)$	Melodic coda licensing
M5	$\mathbf{fake} \rightarrow O \wedge \bigwedge_{m \neq \mathbf{fake}} \neg m$	Fake onsets

Remember that GP allows languages to impose further restrictions on the melody by recourse to licensing constraints. It is easy to see that licensing constraints operating on single PEs can be captured by propositional formulas. The licensing constraint “A must be head”, for instance, corresponds to the propositional formula $\neg \bar{A}$. Licensing constraints that extend beyond a single segment can be modeled using \triangleleft and \triangleright , provided their domain of application is finitely bounded (see the discussion on spreading below for further details). Thus licensing constraints pose no obstacle to formalization in our logic, either.

As mentioned above, I use μ to mark “mute” segments that will be realized as the empty string. The distribution of μ is simple for O and C — the latter never allows it, and the former only if it is unary branching and followed by a pronounced N. For N, on the other hand, we first need to distribute \checkmark in a principled manner across the string to mark the licensed nuclei, i.e. those N that may remain unpronounced. Note that unpronounced segments may not contain any other elements (which would affect spreading).

L1	$\mu \rightarrow \bigwedge_{m \notin \{\mu, \checkmark\}} \neg m \wedge \neg C \wedge (N \rightarrow \checkmark)$	Empty categories
L2	$N \wedge \triangleleft N \rightarrow (\mu \leftrightarrow \triangleleft \mu)$	No partially mute branching nuclei
L3	$O \wedge \mu \rightarrow \neg \triangleleft O \wedge \triangleright (N \wedge \neg \mu)$	Mute onsets

$$\begin{array}{c}
\mathbf{L4} \quad N \wedge \checkmark \leftrightarrow \underbrace{\triangleright (C \wedge \bigvee_{i \in \mathcal{S}} i)}_{\text{Magic Licensing}} \vee \underbrace{(\neg \triangleleft N \wedge \blacktriangleright \perp)}_{\text{FEN}} \vee \text{P-licensing} \\
\quad \quad \quad \underbrace{((\neg \triangleleft N \rightarrow \triangleleft (\triangleleft N \vee \blacktriangleleft \perp)) \wedge (\neg \triangleright N \rightarrow \triangleright \triangleright (N \wedge \neg \mu)))}_{\text{Proper Government}}
\end{array}$$

Axiom **L4** looks daunting at first, but it is easy to unravel. The magic licensing conditions tells us that N is licensed if it is followed by a sibilant in coda position.² The FEN condition ensures that wordfinal N are licensed if they are non-branching. The proper government condition is the most complex one, though it is actually simpler than the original GP definition. Remember that N is properly governed if the first N following it is pronounced and neither a branching onset nor a coda intervenes. Also keep in mind that we treat a binary branching constituent as two adjacent unary branching constituents. The proper government condition then enforces a structural requirement such that N (or the first N if we are talking about two adjacent N) may not be preceded by two constituents that are not N and (the second N) may not be followed by two constituents that are not N or not pronounced. Together with axioms **S1-S7**, this gives the same results as the original constraint.³

The last module, spreading, is also the most difficult to accommodate. Most properties of spreading are language specific — only the set of spreadable features and the ban against onset internal spreading are universal. To capture this variability, I define a general spreading scheme σ with six parameters $i, j, \omega, \varpi, min$ and max .

$$\sigma := \bigwedge_{\pi_1(i)=\pi_1(j)} (i \wedge \omega \rightarrow \bigvee_{n=min}^{max} \diamond^n(j \wedge \varpi) \wedge (O \wedge \diamond O \rightarrow \bigvee_{n=min+1}^{max} \diamond^n(j \wedge \varpi)))$$

The variables $i, j \in \mathcal{E}$, coupled with judicious use of the formulas ω and ϖ regulate the optionality of spreading. If spreading is optional, i is a spread element and ω, ϖ are formulas describing, respectively, the structural configuration of the target of spreading and the set of licit sources for spreading operations to said target. If

² Note that we can easily restrict the context, if this appears to be necessary for empirical reasons. Strengthening the condition to $\triangleright (C \wedge \bigvee_{i \in \mathcal{S}} i) \wedge \triangleleft \blacktriangleleft \perp$, for example, restricts magic licensing to the N occupying the second position in the string.

³ In this case, the modal logic is once again flexible enough to accommodate various alternatives. For instance, if proper government should be limited to non-branching Ns, one only has to replace both occurrences of \rightarrow by \wedge . Also, my formalization establishes no requirement for a segment to remain silent, because N often are pronounced in magic licensing configurations or at the end of a word in a FEN language. For proper government, however, it is sometimes assumed that licensed nuclei have to remain silent, giving rise to a strictly alternating pattern of realized and unrealized Ns. If we seek to accommodate such a system, we have to distinguish Ns that are magically licensed or FEN licensed from Ns that are licensed by virtue of being properly governed. The easiest way to do so is to split \checkmark into two features \checkmark_o and \checkmark_m (optional and mandatory), the latter of which is reserved for properly governed Ns. The simple formula $\checkmark_m \rightarrow \mu$ will force such Ns to remain unpronounced.

spreading is mandatory, then i is a local element and ω , σ describe the source and the set of targets. If we want spreading to be mandatory in only those cases where a target is actually available, ω has to contain the subformula $\bigvee_{n=min}^{max} \diamond^n \sigma$. Observe moreover that we need to make sure that every structural configuration is covered by some ω , so that unwanted spreading can be blocked by making σ not satisfiable. As further parameters, the finite values $min, max > 0$ encode the minimum and maximum distance of spreading, respectively. Finally, the operator $\diamond \in \{\triangleleft, \triangleright\}$ fixes the direction of spreading for the entire formula (\diamond^n is the n -fold iteration of \diamond). With optional spreading, the direction of the operator is opposite to the direction of spreading, otherwise they are identical. The different ways of interaction between the parameters is summarized in Table 1.

Mode	Direction	i	ω	σ	\diamond
optional	left	spread	target	source	\triangleright
optional	right	spread	target	source	\triangleleft
mandatory	left	local	source	target	\triangleleft
mandatory	right	local	source	target	\triangleright

Table 1. Parameterization of spreading patterns with respect to σ

As the astute reader (or rather, all readers that took a glimpse at footnotes 2 and 3) will have noticed by now, nothing in our logic prevents us from defining alternative versions of GP. Whether this is a welcome state of affairs is a matter of perspective. On the one hand, the flexibility of our logic ensures its applicability to a wide range of different variants of GP, e.g. to versions where spreading is allowed within onsets or where the details of proper government and the restrictions on branching vary. On the other hand, it raises the question whether there isn't an even weaker modal logic that is still expressive enough to formalize GP. However, the basic feature calculus of GP already requires the logical symbols \neg and \wedge , which gives us the complete set of logical connectives, and we furthermore need \triangleleft and \triangleright to move us along the phonological string. Hence, imposing any further syntactic restrictions on formulas requires advanced technical concepts such as the number of quantifier alternations. But this brings us back to an issue I discussed in the preface to this section: the loose grip of mathematical methods, and why it isn't as problematic as it might seem initially. Lest I unnecessarily bore the reader with methodological remarks, I shall merely point out that it is doubtful that a further weakening of the logic would have interesting ramifications given the questions I set out to answer; I am not interested in the logic that provides the best fit for a specific theory but in the investigation of entire classes of string-based phonological theories from a model-theoretic perspective. In the next section, I try to get closer to this goal.

2 The Parameters of Phonological Theories

2.1 Elaborate Spreading — Increasing the Generative Capacity

It is easy to see that the modal logic defined in the previous section is powerful enough to account for all finitely bounded phonological phenomena (I hasten to add that this does not imply that GP itself can account for all of them, since certain phenomena might be ruled out by, say, the syllable template or the ECP). In fact, it is even possible to accommodate many long-distance phenomena in a straight-forward way, provided that they can be reinterpreted as arising from iterated application of finitely bounded processes or conditions. Consider for example a stress rule for language L that assigns primary stress to the last syllable that is preceded by an even number of syllables. Assume furthermore that secondary stress in L is trochaic, that is to say it falls on every odd syllable but the last one. Let 1 and 2 stand for primary and secondary stress, respectively. Unstressed syllables are assigned the feature 0. Then the following formula will ensure the correct assignment of primary stress, even though the notion of being separated from the left word edge by an even number of syllables is unbounded (for the sake of simplicity, I assume that every node in the string represents a syllable; it is an easy but unenlightening exercise to rewrite the formula for a GP syllable template consisting of Os, Ns and Cs).

$$\bigvee_{i \in \{0,1,2\}} i \wedge \bigwedge_{i \neq j \in \{0,1,2\}} (i \rightarrow \neg j) \wedge (\blacktriangleleft \perp \rightarrow 1 \vee 2) \wedge (2 \rightarrow \blacktriangleright 0) \wedge \\ (0 \rightarrow \blacktriangleright (1 \vee 2) \vee \blacktriangleright \perp) \wedge (1 \rightarrow \neg \blacktriangleleft 1 \wedge (\blacktriangleright \perp \vee \blacktriangleright \perp))$$

Other seemingly unbounded phenomena arising from iteration of local processes, most importantly vowel harmony (see [3] for a GP analysis), can be captured in a similar way. However, there are several unbounded phonological phenomena that require increased expressivity, as I discuss *en detail* in [10].

Since we are only concerned with string structures, it is a natural move to try to enhance our language with operators from more powerful string logics, in particular, linear temporal logic. The first step is the addition of two operators \blacktriangleleft^+ and \blacktriangleright^+ with the corresponding relation R_{\blacktriangleleft^+} , the transitive closure of R_{\blacktriangleleft} . This new logic is exactly as powerful as restricted temporal logic [6], which in turn has been shown to exactly match the expressivity of the two-variable fragment of first-order logic ([7]; see [44] for further equivalence results). Among other things, unbounded OCP effects [9, 26] can now be captured in an elegant way. The formula $O \wedge \underline{A} \wedge \bar{L} \wedge \bar{T} \rightarrow \blacktriangleright^+ \neg(O \wedge \underline{A} \wedge \bar{T})$, for example, disallows alveolar nasals to be followed by another alveolar stop, no matter how far the two are apart.

But \blacktriangleleft^+ and \blacktriangleright^+ are too coarse for faithful renditions of unbounded spreading. For example, it is not possible to define all intervals of arbitrary size within which a certain condition has to hold (e.g. no b may appear between a and c). As a remedy, we can add to the logic the until and since operators U and S familiar from linear temporal logic, granting us the power of full first-order logic and

pushing us to the level of the star-free languages [5, 6, 29, 41]. Star-free languages feature a plethora of properties that make them very attractive for purposes of natural language processing. Moreover, the only phenomenon known to the author that exceeds their confines is stress assignment in Cairene Arabic and Creek, which basically works like the stress assignment system outlined above — with the one exception that secondary stress is not marked overtly [12, 30]. Under these conditions, assigning primary stress involves counting modulo 2, which is undefinable in first-order logic, whence a more powerful logic is needed. The next step up from the star-free stringsets are the regular stringsets, which can count modulo n . The regular stringsets are identical to the sets of finite strings definable in monadic second order logic (MSO) [2], linear temporal logic with modal fixed point operators [43] or regular linear temporal logic [27]. In linguistic terms, this corresponds to spreading being capable of picking its target based on more elaborate patterns, counting modulo 2 being one of them. For further discussion of the relation between expressivity and phenomena in natural language phonology, the reader is once again referred to [10].

A caveat is in order, though. Thatcher [40] proved that every recognizable set is a projection of some local set. Thus the hierarchy outlined above collapses if we grant ourselves an arbitrary number of additional features to encode all the structural properties our logic cannot express. In the case of primary stress in Cairene Arabic and Creek, for instance, we could just use the feature for secondary stress assignment even though secondary stress seems to be absent in these languages. Generally speaking, we can reinterpret any unbounded dependency as a result of iterated local processes by using “invisible” features. Therefore, all claims about generative capacity hold only under the proviso that all such coding-features are being eschewed.

We have just seen that the power of GP can be extended along the subregular hierarchy, up to the power of regular languages, and that there seems to be empirical motivation to do so. Interestingly, it has been observed that SPE yields regular languages, too [15, 17]. But even the most powerful rendition of GP defines only a proper subset of the stringsets derivable in SPE, apparently due to its restrictions on the feature system, the syllable template and its government requirements. The question we face, then, is whether we can generalize GP in these regards, too, to push it to the full power of SPE and obtain a multidimensional vector space of phonological theories.

2.2 Feature Systems

It is easy to see that at the level of classes of theories, the restriction to privative features is immaterial. A set of PEs is denoted by some propositional formula over \mathcal{E} , and the boolean closure of \mathcal{E} is isomorphic to $\wp(\mathcal{E})$. But as shown in [22], a binary feature system using a set of features \mathcal{F} can be modeled by the powerset algebra $\wp(\mathcal{F})$, too. So if $|\mathcal{E}| = |\mathcal{F}|$, then $\wp(\mathcal{E})$ and $\wp(\mathcal{F})$ are isomorphic, and so are the two feature systems. The same result holds for systems using more than two feature values, provided their number is finitely bounded, since multivalued features can be replaced by a collection of binary valued features given sufficient

co-occurrence restrictions on feature values (which can easily be formalized in propositional logic).

One might argue, though, that the core restriction of privative feature systems does not arise from the feature system itself but from the methodological principle that absent features, i.e. negative feature values, behave like constituency information and cannot spread. In general, though, this is not a substantial restriction either, as for every privative feature system \mathcal{E} we can easily design a privative feature system $\mathcal{F} := \{e^+, e^- \mid e \in \mathcal{E}\}$ such that $\mathfrak{M}, w \models e^+$ iff $\mathfrak{M}, w \models e$ and $\mathfrak{M}, w \models e^-$ iff $\mathfrak{M}, w \models \neg e$. Crucially, though, this does not entail that the methodological principle described above has no impact on expressivity when the set of features is fixed across all theories, which is an interesting issue for future research.

2.3 Syllable Template

While GP’s syllable template could in principle be generalized to arbitrary numbers and sizes of constituents, a look at competing theories such as SPE and CVCV [28, 38] shows that the number of different constituents is already more than sufficient. This is hardly surprising, because GP’s syllable template is modeled after the canonical syllable template, which isn’t commonly considered to be in need of further refinement. Consequently, we only need to lift the restriction on the branching factor and allow theories not to use all three constituent types. SPE then operates with a single N constituent of unbounded size (as no segment in SPE requires special licensing, just like Ns in GP), whereas CVCV uses N and O constituents of size 1. Regarding the government relations, the idea is to let every theory fix the branching factor b for each constituent and the maximum number l of licensees per head. Every node within some constituent has to be constituent licensed by the *head*, i.e. the leftmost node of said constituent. Similarly, all nodes in a coda or non-head position have to be transconstituent licensed by the head of the following constituent. For every head the number of constituent licensees and transconstituent licensees, taken together, may not exceed l .

Even from this basic sketch it should already be clear that the syllable template can have a negative impact on expressivity, but only under the right conditions. For instance, if our feature system is set up in a way such that every symbol of our alphabet is to be represented by a PE in N (as happens to be the case for SPE), restrictions on b and l are without effect. Thus one of the next stages in this project will revolve around determining under which conditions the syllable template has a monotonic effect on generative capacity.

2.4 Representations versus Derivations

One of the most striking differences between phonological theories is the distinction between representational and derivational ones, which begs the question how we can ensure comparability between these two classes. Representational theories are naturally captured by the declarative, model-theoretic approach,

whereas derivational theories like SPE are usually formalized as regular relations [17, 31], which resist being recast in logical terms due to their closure properties. This problem is aggravated by the fact Optimality Theory [33], which provides the predominant framework in contemporary phonology, is also best understood in terms of regular relations [8, 16]. Of course, one can use a coding trick from two-level phonology [23] and use an unpronounced feature like μ to ensure that all derivationally related strings have the same length, so that the regular relations can be interpreted as languages over pairs and hence cast in MSO terms [42]. Unfortunately, it is far from obvious how this method could be extended to subregular grammars, because Thatcher’s theorem tells us that the projection of a subregular language of pairs might be a regular language. But due to the ubiquity of SPE and OT analyses in phonology, no other open issue is of greater importance to the success of this project.

3 Conclusion

The purpose of this paper was to lay the foundation for a general framework in which string-based phonological theories can be matched against each other. I started out with a modal logic which despite its restrictions was still perfectly capable of defining a rather advanced and intricate phonological theory. I then tried to generalize the theory along several axes, some of which readily lent themselves to conclusive results while others didn’t. We saw that the power of spreading, by virtue of being an indicator of the necessary power of the description language, has an immediate and monotonic effect on generative capacity. Feature systems, on the other hand, were shown to be a negligible factor in theory comparisons; it remains an open question if the privativity assumption might affect generative capacity when the set of features is fixed. A detailed study of the effects of the syllable template also had to be deferred to later work. Clearly the most pressing issue, though, is the translation from representational to derivational theories. Not only will it enable us to reconcile two supposedly orthogonal perspectives on phonology, but it also allows us to harvest results on finite-state OT [8] to extend the framework to optimality theory. Even though a lot of work remains to be done and not all of my goals may turn out be achievable, I am confident that a model-theoretic approach provides an interesting new perspective on long-standing issues in phonology.

References

- [1] Blackburn, P., de Rijke, M., Venema, Y.: *Modal Logic*. Cambridge University Press, Cambridge (2002)
- [2] Büchi, J.R.: Weak second-order arithmetic and finite automata. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 6, 66–92 (1960)
- [3] Charette, M., Göksel, A.: Licensing constraints and vowel harmony in Turkic languages. *SOAS Working Papers In Linguistics and Phonetics* 6, 1–25 (1996)

- [4] Chomsky, N.: Lectures on Government and Binding: The Pisa Lectures. Foris, Dordrecht (1981)
- [5] Cohen, J.: On the expressive power of temporal logic for infinite words. *Theoretical Computer Science* 83, 301–312 (1991)
- [6] Cohen, J., Perrin, D., Pin, J.E.: On the expressive power of temporal logic. *Journal of Computer and System Sciences* 46, 271–294 (1993)
- [7] Etessami, K., Vardi, M.Y., Wilke, T.: First-order logic with two variables and unary temporal logic. In: *Proceedings of the 12th Annual IEEE Symposium on Logic in Computer Science*. pp. 228–235 (1997)
- [8] Frank, R., Satta, G.: Optimality theory and the generative complexity of constraint violability. *Computational Linguistics* 24, 307–315 (1998)
- [9] Goldsmith, J.: *Autosegmental Phonology*. Ph.D. thesis, MIT (1976)
- [10] Graf, T.: Comparing incomparable frameworks: A model theoretic approach to phonology. In: *University of Pennsylvania Working Papers in Linguistics*. vol. 16, p. Article 10 (2010), available at: <http://repository.upenn.edu/pwpl/vol16/iss1/10>
- [11] Graf, T.: *Logics of Phonological Reasoning*. Master’s thesis, University of California, Los Angeles (2010)
- [12] Haas, M.R.: Tonal accent in Creek. In: Hyman, L.M. (ed.) *Southern California Occasional Papers in Linguistics*, vol. 4, pp. 195–208. University of Southern California, Los Angeles (1977), reprinted in [39]
- [13] Harris, J., Lindsey, G.: The elements of phonological representation. In: Durand, J., Katamba, F. (eds.) *Frontiers of Phonology*, pp. 34–79. Longman, Harlow, Essex (1995)
- [14] Jensen, S.: Is ? an element? Towards a non-segmental phonology. *SOAS Working Papers In Linguistics and Phonetics* 4, 71–78 (1994)
- [15] Johnson, C.D.: *Formal Aspects of Phonological Description*. Mouton, The Hague (1972)
- [16] Jäger, G.: Gradient constraints in finite state OT: The unidirectional and the bidirectional case. In: Kaufmann, I., Stiebels, B. (eds.) *More than Words. A Festschrift for Dieter Wunderlich*, pp. 299–325. Akademie Verlag, Berlin (2002)
- [17] Kaplan, R.M., Kay, M.: Regular models of phonological rule systems. *Computational Linguistics* 20(3), 331–378 (1994)
- [18] Kaye, J.: Government in phonology: the case of Moroccan Arabic. *The Linguistic Review* 6, 131–159 (1990)
- [19] Kaye, J.: Do you believe in magic? The story of s+C sequences. *Working Papers in Linguistics and Phonetics* 2, 293–313 (1992)
- [20] Kaye, J.: A user’s guide to government phonology (2000), <http://134.59.31.7/scheer/scan/Kaye00guideGP.pdf>, unpublished manuscript
- [21] Kaye, J., Lowenstamm, J., Vergnaud, J.R.: Constituent structure and government in phonology. *Phonology Yearbook* 7, 193–231 (1990)
- [22] Keenan, E.: *Mathematical structures in language* (2008), ms., University of California, Los Angeles

- [23] Koskenniemi, K.: Two-level morphology: A general computational model for word-form recognition and production. Publication 11 (1983)
- [24] Kracht, M.: Syntactic codes and grammar refinement. *Journal of Logic, Language and Information* 4, 41–60 (1995)
- [25] Kracht, M.: Features in phonological theory. In: Löwe, B., Malzkorn, W., Räsch, T. (eds.) *Foundations of the Formal Sciences II, Applications of Mathematical Logic in Philosophy and Linguistics*, pp. 123–149. No. 17 in *Trends in Logic*, Kluwer, Dordrecht (2003), papers of a conference held in Bonn, November 11–13, 2000
- [26] Leben, W.: *Suprasegmental Phonology*. Ph.D. thesis, MIT (1973)
- [27] Leucker, M., Sánchez, C.: Regular linear temporal logic. In: *Proceedings of The 4th International Colloquium on Theoretical Aspects of Computing (ICTAC'07)*. pp. 291–305. No. 4711 in *Lecture Notes in Computer Science* (2005)
- [28] Lowenstamm, J.: CV as the only syllable type. In: Durand, J., Laks, B. (eds.) *Current Trends in Phonology: Models and Methods*. pp. 419–421. European Studies Research Institute, University of Salford (1996)
- [29] McNaughton, R., Pappert, S.: *Counter-Free Automata*. MIT Press, Cambridge, Mass. (1971)
- [30] Mitchell, T.F.: Prominence and syllabification in Arabic. *Bulletin of the School of Oriental and African Studies* 23(2), 369–389 (1960)
- [31] Mohri, M., Sproat, R.: An efficient compiler for weighted rewrite rules. In: *In 34th Annual Meeting of the Association for Computational Linguistics*. pp. 231–238 (1996)
- [32] Potts, C., Pullum, G.K.: Model theory and the content of OT constraints. *Phonology* 19(4), 361–393 (2002)
- [33] Prince, A., Smolensky, P.: *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell, Oxford (2004)
- [34] Pullum, G.K.: The evolution of model-theoretic frameworks in linguistics. In: Rogers, J., Kepsers, S. (eds.) *Model-Theoretic Syntax @ 10*. pp. 1–10 (2007)
- [35] Rogers, J.: A model-theoretic framework for theories of syntax. In: *Proceedings of the 34th Annual Meeting of the ACL*. pp. 10–16. Santa Cruz, USA (1996)
- [36] Rogers, J.: Strict LT_2 : Regular :: Local : Recognizable. In: Retoré, C. (ed.) *Logical Aspects of Computational Linguistics: First International Conference, LACL '96 (Selected Papers)*. *Lecture Notes in Computer Science/Lecture Notes in Artificial Intelligence*, vol. 1328, pp. 366–385. Springer (1997)
- [37] Russell, K.: *A Constraint-Based Approach to Phonology*. Ph.D. thesis, University of Southern California (1993)
- [38] Scheer, T.: *A Lateral Theory of Phonology: What is CVCV and Why Should it be?* Mouton de Gruyter, Berlin (2004)
- [39] Sturtevant, W.C. (ed.): *A Creek Source Book*. Garland, New York (1987)
- [40] Thatcher, J.W.: Characterizing derivation trees for context-free grammars through a generalization of finite automata theory. *Journal of Computer and System Sciences* 1, 317–322 (1967)

- [41] Thomas, W.: Star-free regular sets of ω -sequences. *Information and Control* 42, 148–156 (1979)
- [42] Vaillette, N.: Logical specification of regular relations for NLP. *Natural Language Engineering* 9(1), 65–85 (2003)
- [43] Vardi, M.Y.: A temporal fixpoint calculus. In: *Proceedings of the 15th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*. pp. 250–259 (1988)
- [44] Weil, P.: Algebraic recognizability of languages. In: Fiala, J., Koubek, V., Kratochvíl, J. (eds.) *Mathematical Foundations of Computer Science 2004*, *Lecture Notes in Computer Science*, vol. 3153, pp. 149–175. Springer, Berlin (2004)