THE JOURNAL OF THE
UNDERGRADUATE LINGUISTICS
ASSOCIATION OF BRITAIN

ISSN 2754-0820

COMPLEXITY AND THE PHONOLOGICAL TURING MACHINE

Louis Van Steene

University of Cambridge
louis@lvs.me.uk

Keywords:	Natural Phonology; Optimality Theory; computational complexity; Turing machine; phonological process; I-language
JoULAB URL:	https://www.ulab.org.uk/journal/volumes/2/issues/2/articles/1
DOI:	https://doi.org/10.5281/zenodo.8258548
Citation:	Van Steene, L. (2024). Complexity and the Phonological Turing Machine. *Journal of the Undergraduate Linguistics Association of Britain*, 2(2), pp. 11-32.

Received: 30/07/2020
Revisions: 10/02/2022
Accepted: 06/07/2023

JOURNAL OF THE UNDERGRADUATE LINGUISTICS ASSOCIATION OF BRITAIN

Complexity and the Phonological Turing Machine

Louis Van Steene
University of Cambridge

Abstract. Any scientific theory, including within linguistics, requires a coherent philosophical basis in order to evaluate and decode the relations between data, phenomena, and theory. Chomsky's (1964; 1965) identification of explanatory adequacy as one of the ultimate goals of a theory of I-language was a seminal step in this regard. This paper explores this term and its implications with special reference to phonological theory. In particular, four different groups of theories are evaluated on the metric of explanatory adequacy: Chomsky & Halle's (1968) Rule-Based Phonology, Stampe's (1979) Natural Phonology, Optimality Theory (Prince & Smolensky, 1993; McCarthy & Prince, 1993), and Substance-Free Phonology (SFP; Hale & Reiss, 2008; Samuels, 2009). SFP is found to be the most promising from a minimalist, computational perspective. This theoretical foundation is subsequently adapted into Watumull's (2012; 2015) Turing programme for linguistic theory. From this perspective, the mind is viewed as equivalent to a Turing machine – the universal model of computation. Framed as such, a novel approach emerges regarding computational complexity and economy in (phonological) derivations, which are issues otherwise often only nebulously invoked in the evaluation of theories. A method of analysis is introduced by adopting 'Big-O notation' as used for asymptotic analysis in computer science. This method is shown to highlight the importance of strongly defining the inventory of computational primitives and procedures within a theory. Specific suggestions regarding the nature of the most optimal theory under this analysis are made with respect to Samuels' (2009) explicitly Minimalist brand of SFP.

Plain English Abstract. The goal of a particular linguistic theory should be to make falsifiable predictions about what every human (excepting language disorders) knows about language and how they come to have this knowledge. Chomsky (1964) respectively labels these two goals as descriptive adequacy and explanatory adequacy. An important area of linguistic knowledge concerns phonology, the mental combination and manipulation of discrete, arbitrary symbols that are composed into words. These symbols are known as phonemes, for example /p/, /a/, /u/. In this paper, four theories of phonology are presented and evaluated against the goal of explanatory adequacy. This leads to Watumull's (2012) proposal that the various aspects of linguistic knowledge, including the phonological module, be considered as equivalent to a Turing machine, a device capable of performing any computable operation. In such an approach, it becomes apparent that consideration of the precise nature of the inventory of procedures available within a theory is crucial. It is further proposed that theories that take representations as prior to these procedures end up being more complex computationally. In this respect, something can be learned from the study of computational complexity within computer science, in particular the use of 'Big-O notation'. From all this, one theoretical contender emerges as most promising – namely, a version of Samuels' (2009) theory. It is clear, however, that the theory needs to be refined much further in line with these goals.

Keywords: Natural Phonology; Optimality Theory; computational complexity; Turing machine; phonological process; I-language

1 Introduction

When discussing the adequacy of theories, it is important to dissect exactly what one means by the term 'theory' in this context. One crucial element, following the Popperian tradition, is that a scientific theory must stipulate the evidence needed to falsify it (Thornton, 2019); this extends to linguistic theory. When considering exactly what 'evidence' is, difficulty can arise in distinguishing the three interactive levels of *data*, *phenomena*, and *theory*. Bogen & Woodward (1977) claim that 'phenomena' are separate both from data on the one hand, and theory on the other. A theory explains or predicts phenomena, and data are evidence for the existence of phenomena. To complicate matters, the theory itself can influence (the collection of) the data which then inform the postulation of phenomena, which is an issue well-known in the philosophy of science (Kelly, 2016). Thus, there arises a tension here, between accounting for the data directly with the theory on the one hand, and accounting for the phenomena as they actually are on the other. This can be seen as a paraphrasing of the difference between descriptive adequacy, viz. making predictive, testable generalisations that account for the data, and explanatory adequacy, which requires the theory to account for the phenomenon of its acquisition (Chomsky, 1965).

In generative linguistics, the object of study is *I-language*, characterised as *internal, intensional* (generable intensionally, as opposed to being explicit and extensional) and *individual.* This is what gives language its essential creative aspect, corresponding to what Humboldt called the 'infinite use of finite means', discrete infinity – the 'finite means' in question is the I-language[1]. The goal is to achieve a formal description of I-language which meets explanatory adequacy. Simplicity is the goal of any formal, scientific description, with the assumption being that the simplest theory is most likely to be right – and even if this ends up not being the case, such a theory would prove the best foundation for a reality that turns out to be more complex (Hale & Reiss, 2008). This is taken to its ultimate conclusion in the Minimalist programme (Chomsky, 1995), which approaches UG 'from below' in order to establish how little needs to be postulated to specify I-language sufficiently (see Chomsky, 2007). Formulating a metric for this end goal of simplicity, with specific reference to the phonological component of I-language, is the purpose of this paper.

The tension between description and explanation is the topic of Section 2, where the significance of different types of evidence in different theories is explored. In the framework I adopt, I argue that it is not a given that data from phonological acquisition can be taken as evidence for a phonological theory – if anything, any valid evidence that arises should be taken as an unexpected bonus, since the rapidly developing psychology of the child makes controlled delimiting of what is 'phonological' exceptionally difficult. Nonetheless, I conclude that, if treated cautiously, consideration of phonological acquisition is not only useful, but arguably essential to crafting an explanatory theory. In particular, *learnability* constraints are crucial in the domain of competence, whereas 'third factor' considerations like markedness are not directly required in the formal specification of theories of I-language[2]. Following from this conclusion, in Section 3 I expound further on a specific theory of phonological architecture which builds on this epistemological foundation, centring around the Turing machine, an abstract model of computation (Turing, 1936). This is a direct expansion of Watumull's (2012; 2015) Turing programme for linguistics, with expanded application to phonological computation, following the brief outline of Vaux & Watumull (2012). With this foundation in place, I look at a practical and novel application of the Turing computational framework, namely around the idea of 'economy', an oft-used but little-explained proposed constraint on language systems. By adopting measures of computational complexity typically reserved for the domain of computer science, various theories of phonology are evaluated within this novel framework.

---

[1] Note that Humboldt's aphorism conflates *knowledge* of I-language (competence) with *use* of language (performance). This sidesteps the Aristotelian distinction between knowledge and use, as noted variously by Chomsky (see Chomsky 1986; 1995, p. 14; 2006, pp. ix, 15, 113; 2014, p. 4; 2021).

[2] Using the term 'third factors' in the sense of Chomsky (2005, p. 6): "[p]rinciples not specific to the faculty of language", as opposed to the genetic endowment (first factor) and Primary Linguistic Data (the input; second factor). Whilst expressed most clearly by Chomsky (2005), it is worth noting that the interaction of the three factors is found in any biological system (Lewontin, 2000), noted for language at least as early as Chomsky (1965, p. 59).

Received: 30/07/2020
Revisions: 10/02/2022
Accepted: 06/07/2023

2 Phonological Theory

Any theory of phonology attempts to resolve the tensions between competence and performance, description and explanation. This ultimately comes down to two fundamental questions, for which I set out to find an answer in this section: (i) what is a theory of phonology? and (ii) what makes said theory ‘explanatory’? The main focus will be on the developments relating Chomsky & Halle’s (1968) *SPE*-style rule-based phonology (RBP); Stampe’s (1979) Natural Phonology (NP); Optimality Theory (OT) as arbitrated by Prince & Smolensky (1993), McCarthy & Prince (1993), *inter alia*[3]; and Hale & Reiss’s (2008) Substance-Free Phonology (SFP).

To begin with, it is worth elaborating further upon the tension identified in the introduction between description and explanation: what would a non-explanatory yet perfectly descriptive theory, or analysis of a set of data, look like? As a foil, one might first identify what an analysis that is neither explanatory nor descriptive would look like. One could conceive of an ‘analysis’ that merely restates the data, quite literally, without any level of abstraction. This corresponds the most basic level of theoretical adequacy identified by Chomsky (1964, p. 28f), *observational* adequacy. A theory that meets observational adequacy is hardly a theory at all, and barely worth attention except as a prerequisite for further, descriptive analysis. Chomsky (1964, fn. 1) does note, however, that this step is not entirely trivial: ‘the fact that a certain noise was produced, even intentionally, by an English speaker does not guarantee that it is a well-formed specimen of his language’. ‘[D]eviant utterances’, which may be entirely appropriate to the circumstance, must therefore be distinguished from ‘well-formed specimen[s]’. A significant factor in generating deviance is the volatility of *performance*, as distinguished from *competence* (Chomsky, 1986)[4]. This becomes particularly acute in the case of child language and as associated with learnability considerations, as will become evident in subsequent discussion.

Assuming that the level of observational adequacy can be met, the task is then to specify this data ‘in terms of significant generalizations that express underlying regularities in the language’ (Chomsky, 1964, p. 28); in other words, to provide a descriptive account. Two versions of a descriptive account for an elementary phenomenon of English phonology, the realisation of plural /-z/, are presented in (1) (courtesy of Vaux, p.c.), abbreviated in typical ways. Both contain generalisations and would predict the same typical observations of the phenomenon in question, and thus both appear to meet descriptive adequacy.

(1) (a) /-z/ → [s] / { p t k … } _
/-z/ → [z] / { b d g … } _
/-z/ → [əz] / { s z ʃ ʒ … } _
(b) Epenthesis: Ø → [ə] / _ <C>#
Voicing Assimilation: [-son] → [αvoice] / [-son, αvoice] _]$_\sigma$

Multiple questions then present themselves, all considered within the realm of *explanation*. They stem from the most fundamental: how does one select amongst extensionally equivalent descriptive theories?

---

[3] For reasons of scope, discussion of extensions to the original OT model such as Stochastic OT (Boersma & Hayes, 2001) is mostly avoided, although see Section 4.

[4] Another factor, more relevant to syntactic deviance, constitutes a major result of the Minimalist programme: deviance can be the result of an expression being well-formed at one syntactic interface but not the other. Confusion at the conceptual-intentional interface leads to nonsense sentences; confusion at the sensorimotor interface leads to ineffability. See Chomsky et al. (2019, p. 238). For a definition of the ‘interfaces’, which play only a limited role in the subsequent discussion, see Chomsky (2004, p. 106).

In turn, how do we account for the actual intuitions of the speaker? One means of narrowing the choices presented in (1) down would be to appeal to so-called *external* evidence (Kenstowicz & Kisseberth, 1979; Ohala, 1986). For example, one could perform an experiment with nonce words with non-English phonemes in word-final position to see how native English speakers choose to pluralise them – an extension of the 'wug test' (Berko, 1958). Another major consideration is that of *learnability*: a theory consisting of a large number of overly specific rules is evidently less learnable than a system in which the rules hang together in some reasonable way. One potential way of quantifying this is Minimum Description Length, viz. the fewer rules needed the better, a common metric within generative phonology, as will be discussed (cf. Kenstowicz, 1994). With all this in mind, (1b) appears to be the more explanatory account. Lastly, another consideration, brought to light in particular in the context of the Minimalist programme, is that of *evolvability*: UG should be as minimal as possible such that its rapid evolution appears plausible; in other words, the theoretical devices we are assuming to be instantiated biologically should be as limited as is viable. As Chomsky (2007) terms it, we should approach UG 'from below'. Hence, the tension becomes apparent: with a descriptive account, the goal is to account for as large a proportion of the facts as possible; with an explanatory theory, a major driving force is to *reduce* the burden of this account to make it viable in the contexts of acquisition and evolution. An understanding of this tension proves critical in the evaluation of different theories that follows.

With this in place, I offer a brief overview of the central assumptions of the theories in question. The first eight chapters of Chomsky & Halle's (1968) seminal rule-based analysis constitute the foundational document of the generative approach to phonology, which all the subsequent theories inherit from to at least some extent. The central hallmarks of the theory Chomsky & Halle (1968) present are as follows: (i) the grammar consists of rewrite rules which operate on sequences of phonemes, potentially transforming them into a different sequence; (ii) these rules apply cyclically; (iii) the input to the phonological component is a sequence of base forms known as *underlying representations*, originally stored in the lexicon and possibly manipulated by morphosyntactic processes[5]; (iv) the output of the phonological component is a *surface representation*, representing phonetic instructions of some sort. Of the other theories to be considered, SFP sticks most closely to this model, though with crucial differences that will become apparent. The case with NP is similar, although it introduces greater considerations of *markedness*, a concept that only features in the epilogue to *SPE*. Markedness, which has its roots in the traditional phonological study of the Prague School (Trubetzkoy, 1958 [1969]), is considered, in NP and in the epilogue of *SPE*, to be a potential source of evaluation between extensionally equivalent theories, and thus as a way of meeting explanatory adequacy. As Hume (2011) demonstrates, a universally applicable definition of 'markedness' is extremely difficult to parse from the literature, and the concept is perpetually confused, which makes it immediately problematic as an evaluation metric. Here, a somewhat vague definition, whereby something (a phoneme; a feature; a combination thereof) that is 'marked' is in some way less plausible or optimal than something unmarked, will suffice. As the final theory to be considered, OT makes a somewhat radical departure from assumption (i): the OT grammar does not consist of rewrite rules, but of *ranked constraints*. Instead of these being applied cyclically in sequence, as per assumption (ii), potential surface representations are rated against one another, based on how they fare against these constraints. This results in (depictions of) derivations taking on a unique, tabular format. I present a

---

[5] Leaving aside the question of at exactly which point in the derivation phonological forms are introduced, cf. the Distributed Morphology literature (Halle & Marantz, 1993; Marantz, 1997).

Received: 30/07/2020
Revisions: 10/02/2022
Accepted: 06/07/2023

simple example in (2), from Kager (1999, p. 15f), demonstrating how final devoicing of Dutch /bɛd/ ‘bed’ would be evaluated.

(2)

Candidates	*VOICED-CODA	IDENT-IO(voice)
a. ☞ [bɛt]		*
b. [bɛd]	*!	

Key: *VOICED-CODA can be read as ‘voiced obstruent in coda is marked; IDENT-IO(voice) means ‘corresponding input and output segments should match for [±voice]. The candidates are produced by GEN, the workings of which are not considered relevant for the evaluation (but cf. Section 4 below). Asterisks indicate violations of constraints. Candidates are in effect in a race: (b) loses, because *VOICED-CODA outranks IDENT-IO(voice), as indicated by the exclamation mark. See Kager (1999) for a comprehensive introduction to Classic OT.

Markedness considerations are even more tightly entwined with OT, as is evident from the nature of the constraints, and as will be critiqued below.

Returning now to the questions posed at the start of this section, it appears that question (i) (‘what is a theory of phonology?’) is perhaps a bit of a trick, seeing as there are two distinct yet interrelated answers, the result of an ambiguity which is made explicit at least by Chomsky (1965). The dichotomy is one of ontology versus epistemology, which can be paraphrased as ‘theory of a language’ versus ‘theory of language’. In order to be explanatory, a theory of language must explain how the learner acquires their theory of *a* language, the specific language acquired as triggered by the input. The stronger thesis, implied by the systematic ambiguity, is that these are identical (Chomsky, 1986, p. 3; cf. Miller et al., 2016). Acquiring an I-language is clearly a sizeable task for the learner; nevertheless, barring pathology, children acquire their native language with invariable success, rapidly, and on the basis of impoverished input (following the ‘poverty of the stimulus’ argument, Chomsky, 1986, p. 7). As part of their theory, all the theories discussed here include some innate, genetic component to aid the acquisition process, which will be referred to as Universal Grammar (UG), a traditional term adapted by Chomsky to a modern context (see Chomsky, 1964, p. 6; 1966). To tie this down to phonology specifically, it is assumed that phonology is a component of I-language, but as to where to draw the boundary with morphosyntax (as another part of I-language) on the one hand and phonetics (as an extra-grammatical domain) on the other remains initially ambiguous, as these boundary disputes are highly theory-dependent. For instance, RBP and SFP both necessitate a highly modular approach to I-phonology at least with respect to phonetics, especially the latter which separates the two entirely. They can assume the ‘full competence’ or ‘strong identity’ hypothesis (Smith, 1973; Hale & Reiss, 1998), whereby the child and adult are separated not by divergence in competence but only in performance, e.g., differences in production caused by articulatory difficulty due to the incomplete development of the child’s vocal tract. This is as compared with OT, for instance – in a 1988 paper McCarthy, who later became one of the founders of OT, makes his view clear that he sees incorporating phonetics into the phonology as not a redundant but a beneficial endeavour. NP is similarly ‘grounded’ in both phonetics and language acquisition, and OT inherits its approach in considering child phonology and its development into adult phonology evidence of there being ‘natural’ markedness constraints that persist in adult systems. As Stampe, the architect of NP, puts it, what the speaker brings to the language

(Stampe's 'processes') is just as important as what the language brings to the speaker (Stampe's 'rules'). The appropriate OT parallel is between W(ellformedness)-constraints and F(aithfulness)-constraints, respectively – in (2), *VOICED-CODA is a wellformedness constraint and IDENT-IO(voice) is a faithfulness constraint. On the simplest interpretation, W>>F (where >> means 'outranks') in the initial constraint ranking. The raising of faithfulness constraints and the demotion of wellformedness constraints by the learning algorithm leads to marked ('ill-formed') structures eventually being output (Smolensky, 1996).

One (not in fact redundant) fact of language acquisition is thus that language is learnable – in the terms discussed above, our theory of phonology (as a component of I-language) must be explanatory. So, an explanatory theory must be 'grounded' in acquisition in this sense. So-called 'natural' theories like NP and OT descend in altitude by necessitating that I-phonology also account for markedness, in the form of innate phonological processes or constraints brought to the language by the learner, to borrow Stampe's phrasing. Indeed, this also represents a divergence between classic RBP and SFP: in the epilogue of *SPE*, Chomsky & Halle propose that the grammar also outputs language-specific acceptability gradings based on markedness. This parallels a similar idea in Chomsky's syntax at the time: the difference between the 'Core Grammar' and the 'Periphery' as expressed in Chomsky (1981), and the parallel idea in Chomsky (1986) that the output of the grammar is a pair <Output representation, Wellformedness>. In SFP as espoused by Hale & Reiss (2008) and in Samuels' (2009) Minimalist program for phonology, for instance, this formulation is removed from the theory of competence, paralleling developments in syntax (in which such properties arise at the interfaces, cf. Chomsky et al. 2019, p. 238). Instead of being part of the grammatical, computational, phonological module, wellformedness considerations arise from performance systems, including the phonetic component, following Anderson's (1981) line. Two divergent approaches to the facts of phonological acquisition are thus evident, regarding whether or not an explanatory theory of phonology needs to account for surface child phonology phenomena or not – indeed, whether 'child phonology' is phonology at all (cf. Smith, 1973; 2009).

This links to an idea from Hale & Reiss (2008), who frame three 'problems' confronted by the linguist, presented in (1).

(3) (a) **The (weak) AI problem**: to create a model that is 'weakly equivalent' to language – i.e. a model that generates the same output strings as language.
(b) **The Human's problem**: the learner acquires only one grammar[6], rather than any number of other extensionally equivalent grammars.
(c) **The Linguist's problem**: to work out how learners solve the Human's problem.

It is apparent that the Human's problem links up with the ontological question, 'what is the theory of a language?', and the Linguist's problem links up with the epistemological question, 'what is the theory of language?' Everyone has to figure out a theory of their language, whereas only linguists care about *the* theory of language. So where does the AI problem fit into this? In line with Hale & Reiss, I conclude that the AI problem has no bearing on an *explanatory* theory of phonology, and in turn on language acquisition problems generally. Machine learning AI, in the form of deep neural networks that generate

---

[6] This wording to avoid the tautological 'the learner acquires only the grammars they acquire', despite the fact that retaining the latter would continue the ambiguity of my use of 'learner' to refer not just to first, but to bilingual first, and second language acquirers. For relevant comments regarding the endemic focus on the 'ideal' monolingual in much work in theoretical linguistics, cf. Bley-Vroman (2009, p. 180) and, correspondingly, Chomsky (1997, p. 128).

Received: 30/07/2020
Revisions: 10/02/2022
Accepted: 06/07/2023

weakly equivalent output, are not 'intelligent' in a way that resembles human capacity, and they do not meet the criteria for explanation as defined *supra*. On the contrary, the Linguist's problem is to create a theory equivalent to a *strong* AI with genuine human intelligence – the kind Turing envisioned, and that which will be returned to below, in Section 3[7].

This has direct consequences for (classic) OT as a viable theory of I-language. Most strikingly, if a neutral/random ordering of the constraint set CON is innate (part of UG), CON must also be especially large and complex in order to account for phonological phenomena in every studied language, let alone every possible language or even every humanly computable language, the latter of which it is arguably impossible to come to a constraint set for anyway. Furthermore, it seems like an unnecessarily difficult task for the learner to continuously reorder CON as dictated by each parsed input string to the learning algorithm. This is difficult to reconcile with two of the aforementioned crucial facts of language acquisition – its speed and reliance on impoverished input. These two problems in particular are discussed in more detail in Section 4 below. OT's learnability problems and extensional nature are taken by Hale & Reiss (2008) to demonstrate that OT has abandoned the endeavour of solving the Linguist's/Human's problem of finding an intensionally equivalent grammar, and instead can only solve the extensionally equivalent AI problem. In other words, they claim that OT is an inconceivable model of human phonological competence and is merely a point of mathematical interest.

Next, I move onto another set of 'facts' about phonological acquisition, namely empirical data from child production and perception. As explained in Section 1, following the Popperian ideal the use of this data as evidence depends strongly on the theory. Theories like SFP which put acquisition down to grammar-external performance factors make no use of child phonology data. NP and OT may then be able to gain empirical ground on SFP by killing two wugs with one stone[8] – that is, accounting for both child phonology and adult phonology in the same model, as a unified, contingent phenomenon. If this can be done, it may be a sign that SFP is on the wrong track as a theory of I-language. Unfortunately, in OT's case, the fatal theoretical flaws discussed above extend into the empirical domain. A well-attested fact of phonological acquisition is the spontaneous emergence of what Smith (1973) calls 'puzzles' – opaque chain shifts. Smith's example from his case study of his son 'A' is as in (2):

(4) 'puzzle' /pʌzəl/ -> [pʌdəl]
'puddle' /pʌdəl/ -> [pʌgəl]

Clearly A is able to produce the phone [d]; the issue is in the *mapping* from underlying to surface forms. This is trivial to account for in a rule-based system like RBP, adopted by Smith, or SFP (although I leave the issue of why such a rule would arise aside, cf. Bach & Harms, 1972). Classic OT, on the other hand, struggles. There is no constraint ranking which can create this kind of opacity (Hale & Reiss, 2008). This and other unnatural processes found in child language, such as the Click Girl (Bedore et al., 1994) who substituted clicks for consonants, but was able to rapidly unlearn this with instruction, is problematic for theories grounded in acquisition in this sense, and prefers a rule-based analysis over constraints, since rapid unlearning of an opaque rule requires only trivial adjustment. The reasoning behind W>>F and the focus on markedness thus grows increasingly shaky.

---

[7] Note that the 'Turing test', described by Turing (1950), is in reality a poor measure of strong AI, as an extensionally equivalent weak AI would likely perform just as well. The early success of simple computational models that were able to 'pass' the test in certain settings, like Weizenbaum's (1983) ELIZA, is clear evidence of this. Note further Turing's (1950) own opinion, that: "The [] question, 'Can machines think!' [*sic*] I believe to be too meaningless to deserve discussion". See the original Turing (1950) paper, and Watumull (2012) for further discussion.

[8] With apologies for the gruesome image. The word 'wug' was coined by Berko (1958).

NP may be more promising in this regard, being even more explicitly grounded in child phonology, whilst not eschewing the notion of 'rule' entirely. Stampe claims that there are innate phonological processes that must be 'suppressed' by the child in the learning process. Some processes survive into the adult grammar and show up especially in casual (i.e., more 'natural') speech. One example of a process that is rife in both adult and child language is assimilatory nasalisation. Stampe cites a number of diary studies that show this; another example from Ferguson (1986) – [buã] for /pɛn/ – shows nasal assimilation followed by deletion of the final nasal, alongside voicing assimilation and some vowel quality changes which appear in part to be driven be coarticulation or other markedness effects. Other processes suggested by Stampe are flapping and flap-deletion, which show up in my dialect (an idiosyncratic Americanised facsimile of Standard Southern British English) and many American dialects of English, plus syllabification. Furthermore, NP processes apply non-linearly, randomly, sequentially, and iteratively – which allows them to derive a lot of different possible output forms even in adult speech. This can account for the difference between 'formal' speech patterns and casual speech, whilst simultaneously providing a full-competence style model for child phonology. Stampe exemplifies this using the example 'divinity fudge'; I show a partial derivation (leaving aside stop aspiration) of a simpler example 'carnival' in (3), where non-asterisked forms are potential outputs:

(5) /kaɹnɪvɫ̩/
Syllabification: *kaɹ.nɪ.vɫ̩
Flapping: *kaɹ.ɾ̃ɪ.vɫ̩
Nasal assimilation: kãɹ̃.ɾ̃ɪ.vɫ̩
Flap deletion: kãɹ̃.ɪ.vɫ̩
[…vowel assimilation, syllabification, vowel shortening …]
[kãə̯˞.vɫ̩]

One can imagine a child utterance that continues the application of processes, further reducing the output. NP is thus able to derive child utterances using the same processes as present in adult grammar – what varies is the degree of suppression. Opacity effects can arise from rules (although again the ontology is murky), and Kiparsky's feeding and bleeding orders are taken care of simultaneously. Another interesting consequence of this regards L2 acquisition: very briefly, Hungarian learners of L2 English appear to apply a process of final devoicing in their L2, despite it being an active process in neither the L1 nor the L2 (Altenberg & Vago, 1983). These kinds of emergent unmarked processes in interlanguage constitute very productive test cases for NP.

However, trying to account for the fact of variability found in child L1 and adult L2 production in the grammar also leads to problems for NP – not least, the sheer scale of variation found in acquisition data, both within and across individuals. I cited Ferguson's child data for /pɛn/ earlier, but I should note here that [buã] was amongst ten different realisations of supposedly the same word in just a half hour period. Controlling for the fact that all of these data may well have been patched by the observer in making their phonetic transcription masks a potentially even greater degree of variation than is apparent. For Hale & Reiss (2008), this is a strong indication that acquisition-oriented theories like NP are on the wrong track, and that such variation is better attributed to performance constraints. As a final point on this matter, child phonology has in the past mostly been studied in terms of *production*, but only rarely in terms of *perception*. This is an unfortunate lacuna as, assuming that there is only one I-phonology for both perception and production, production-oriented theories like NP may turn out to suffer empirically. The work that has been done in this area, for instance by Jacques Mehler and colleagues (cf. Gervain & Mehler, 2010, for review) certainly suggests that the child has a very early capacity for some kind of phonological processing that does not rely on production systems, which obscures the ontology for NP's 'ease of pronunciation' argument for processes. A further related issue is the stipulation of so-called 'richness of the base', a term coined by OT theorists but an idea that is also made explicit by Stampe (1979, pp. 18ff). The claim is that the actual representations the child constructs do not matter, since thanks to NP processes/OT constraints the grammar will output the same surface representation anyway, e.g. /kæ̃t/ vs /kænt/ for *can't*. This seems to eschew an important part of phonology vital to acquisition – the nature of phonological representations, the importance of which was noted even by McCarthy in 1988: 'if the representations are right, then the rules will follow' (p.84). This is not just a theoretical problem but feeds into an empirical gap too: a crucial finding of Smith's (1973) is that children don't recognise their own utterances played back to them. In other words, they can recognise that their productions do not match their (adult-like) representations – indeed, this 'metalinguistic' knowledge is well attested more broadly (see Gombert, 1997, for review). NP and OT cannot account for this, since they claim that the child's production is a direct output of the grammar, which should be recognised by the grammar as an output of said grammar. The more convincing hypothesis is that something is intervening between the grammar and the actual phonetic realisation – a hypothesis that fits well within SFP, where this 'something' is performance.

SFP's position on this data comes down to the phonology-phonetics boundary dispute alluded to earlier. Following a Minimalist view on language (e.g. Chomsky, 1995; Samuels, 2009) the grammar should be as minimal as necessary: everything that can be done by a grammar-external mechanism which is arguably needed anyway should be done by said mechanism. I-phonology is a computational device that algebraically transforms inputs to outputs which are then transduced into phonetic units at the phonetic interface, to borrow cognitive terminology from Pylyshyn (1984). On Watumull's (2012) interpretation, I-language can be seen as equivalent to the Turing machine. This emphasises the requirement for a highly formal, generative theory of phonology, which NP and OT reject. Substance-free rules can easily be implemented in a phonological Turing machine, as will be elaborated in the following sections. This also constrains the material involved in learnability, be this the requirement for innate features following Hale & Reiss's interpretation of the Subset Principle, or a neo-emergentist approach without innate features *per se* at all, with greater appeal to third factors (Chomsky, 2005; Samuels, 2009; Biberauer, 2019, in syntax). This offers a highly constrained perspective on phonological competence and its acquisition, which OT and NP sacrifice along with any hope of solving the Linguist's/Human's problem.

To sum up this section: language has to be acquired, and so a theory of I-language has to account for its acquisition. (Classic) OT is unable to account for acquisition – rather, it is a theory of 'E-phonology' which achieves surface predictive power at the expense of a plausible ontology. This hinges on the proposed answer to (i): a theory of phonology should be constrained to I-language, i.e. should solve the 'Linguist's problem'. In line with the Chomskyan worldview, the answer to (ii) ('what makes a theory of phonology 'explanatory'?') revolves around the issue of learnability. Learnability considerations should constrain a theory of phonology and are a useful heuristic for evaluation. On the other hand, 'facts of language acquisition' extends to a group of phenomena that can be labelled 'child phonology'. If the child's phonological competence is the same as the adult's, as claimed by Smith and Hale & Reiss, then so-called 'child phonology' is not phonology at all, but rather the result of grammar-external performance factors. Stampe takes the contrary view – that child phonology is a valid object of study, since for child phonology constrains adult phonology, not the other way around. A more modular approach to phonological theory has more potential for explanatory power, on the metrics of falsifiability and predictive power, in the spirit of the scientific programme. Thus, phonological theory should not be ‘grounded’ as many have suggested, but rather should treat acquisition facts as interface considerations in order to reduce the explanatory burden of phonological competence to what is minimally required. Computational complexity is clearly one of these potential burdens, as will be explored in Section 4. With this theoretical foundation, however, it is first time to explore the concept of representation and its counterpart, derivation, in more detail, which will require a description of the phonological Turing machine.

3 The Phonological Turing Machine

I first offer a minimally sufficient definition of the linguistic Turing machine appropriate to the following discussion – for much more detail, see Watumull (2012; 2015).

A Turing machine is a mathematical object that describes a set of computable functions. It consists of three basic components: an unbounded bidirectional tape divided into discrete symbols; a read/write head that can move left and right along the tape; and an instruction table that interprets the symbols that can be read off the tape. During operation, the machine reads the symbol under its head, then the head can move left or right, or write a new symbol onto the tape at its position. This cycle continues indefinitely, or until a ‘halt’ instruction is read. Although apparently simple, Turing proved that this machine could perform any computable operation, given appropriate rules. Since a Turing machine can itself be defined mathematically, it is possible to construct a *universal* Turing machine, which takes the specification of a particular Turing machine’s instruction table as input on the tape. Such programs can consist of any computable operation, given an appropriate ruleset, thus there is no reason to assume that I-language, as a formal system, cannot be implemented on one. A phonological theory of I-language is a stipulation of what the specification, input, and output for the phonological Turing machine should look like, at a certain level of abstraction.

This has consequences for the way we look at phonological operations. To mimic McCarthy: if the derivation is correct, the representation will follow. One can look at the approach detailed above and contrast it with one like McCarthy’s in the following sense: representational theories (such as OT in phonology, perhaps Government & Binding in syntax following Epstein & Seely, 2002) assume that there is an undefined mechanism beyond the scope of the theories that generates candidate representations (cf. GEN in OT). Derivational theories make no comparable assumption, as they inherently must start each derivation from scratch, namely from whatever is taken as primitive. Nevertheless, Brody (2002) argues that it is impossible to escape representations entirely, as a derivational theory remains ‘weakly representational’, because the derivation needs to be able to ‘see into’ previous stages of the derivation, where each ‘stage of the derivation’ is a representation. I disagree, however, that this makes a derivational theory disfavourable, as it is really merely a paraphrasing of the fact that the derivation algorithm needs to be recursive, which has been assumed since the origin of generative grammar (Chomsky, 1957). Indeed, you can still do away with ‘representations’ altogether in a derivational theory, passing the input of the derivation for a particular representation as part of the input for another derivation instead of merely passing said pre-constructed representation. This notational game-playing results in derivation coming on top as epistemically prior to representation in a computational phonology. Once this has been rebutted, it is difficult to see any way of redeeming the representational approach within the formal component, given the specific constraints of explanation detailed above. This is not to depreciate representational approaches entirely – it is by all means useful to look at the output representation of a derivation, for example in an attempt

Received: 30/07/2020
Revisions: 10/02/2022
Accepted: 06/07/2023

to work out what the nature of said derivation is. For instance, representations are a crucial source of data for the phenomenon of derivation, using the terms as defined in Section 1.

Despite this, Ian Roberts (p.c.) does highlight the potential importance of representation at the interface: for syntax, for instance, there has to be some kind of 'representation' at the output, in order that it be interpreted into truth conditions at the conceptual-intentional interface and into a phonological representation at the 'sensorimotor' interface (using terminology from Chomsky, 1995). The same applies to phonology, which effectively transduces a syntactic representation into a phonological one, which is subsequently interpreted by phonetics[9]. Nevertheless, conducive with the derivational approach, another concept can be exapted from computer science, namely that of *streams*. A stream is a serialised flow of information – in computers, bits are transferred more or less one by one – a representation would have to be 'rebuilt' step by step on the other side. This fits naturally into a Turing machine approach, in which effectively only one operation can occur at a time since there is only one read/write head, and it also conveniently parallels the idea of derivation. Information has to be passed around the brain, and if streams are the right answer – and considering their computational simplicity, they may well be – then continually constructing large representations is highly inefficient. This mirrors one of the motivations for phases (Chomsky, 2001), but importantly phases are not a conceptual necessity under the stream approach (see, for instance, the approach taken by Epstein et al., 1998). Empirical arguments in the domain of locality are the more relevant factor in deciding between phases and alternative approaches, which will be discussed little further here (cf. Chomsky, 2008; Rizzi, 2009; Samuels, 2009, for a phonological interpretation). Syntax can be argued to be a Turing machine, phonology can be seen as such, and the same may well go for the semantic-pragmatic component. In spite of the name, Discourse Representation Theory as defined by Kamp & Reyle (1993) is a highly formal, algorithmic theory that would be trivial to implement in a Turing machine. The representations it uses are built and interpreted stepwise from syntactic structure, and are thus perfectly compatible with the information stream approach.

Overall, the approach outlined in this section appears at least to be sufficient to describe I-language and all its putative components. Speculative reasons regarding to what extent it is necessary are the topic of the next section, alongside description of the possible theoretical benefits of its adoption.

4 On Complexity

Adopting this derivational, computational view also allows us to incorporate computational ideas about the nature of complexity. Complexity is an aspect of the phonological component that it is vital to understand, from the perspective both of acquisition and of everyday use of the language faculty. Acquisition and use of language is rapid and based on limited input, and this has led many researchers to adopt principles of 'economy' when deriving phonological theories. Unfortunately, whilst under many mainstream assumptions the most (formally) economical theory is considered the one most likely to be true (Kenstowicz, 1994, i.a.), the idea of economy is often defined weakly, if at all. I suggest that introducing a measure of complexity commonly used in computer science may go some way to resolving this, namely Big-O notation (cf. Knuth, 1997). This is used to represent how the time and space complexity of an algorithm varies with the size of the input(s). An algorithm that completes in constant time $O(1)$ does not vary with the size of the input; one that completes in linear time $O(n)$ varies in direct proportion with the size of the input; etc. Subroutines of different time complexity can

---

[9] Following the architecture proposed by Hale & Reiss (2008). There are alternatives to this worldview, including those that consider phonetics/phonology to be non-distinct (Liberman & Pierrehumbert, 1984; Ohala, 2005), although much of this may be considered definitional (Hale & Reiss, 2000).

be combined, resulting in higher complexities. For instance, a routine of complexity $O(n)$ that occurs n times has the complexity $O(n^2)$. Additionally, a polynomial is always represented simply by the degree of said polynomial (Knuth, 1997). If the aforementioned subroutine also contained a constant-time operation that needed to occur n times, its complexity would be $O(n(n+1)) = O(n^2 + n) = O(n^2)$ – that is, the addition of said operation makes no difference.

There are further levels of time and space complexity that can be described with Big-O beyond $O(1)$, $O(n)$, and $O(n^2)$. The most important options are summarised in Table 1, adapted from Huang (2020), alongside rough evaluations of the expected worst-case performance of an algorithm with such complexity.

Table 1: *Common Big-O complexities*

Notation	**Description**	**Evaluation**
$O(1)$	Constant	Excellent
$O(\log n)$	Logarithmic	Good
$O(n)$	Linear	Fair
$O(n \log n)$	Log-linear	Poor
$O(n^2)$	Polynomial	Very poor
$O(2^n)$	Exponential	Extremely poor
$O(n!)$	Factorial	Extremely poor

Co-opting this approach into phonological theory means is that we do not need to decompose the operations that we take as primitive, assuming that we can estimate the order of magnitude of their time complexity. This is the benefit of the abstraction: we do not necessarily need to deal with neurophysiologically plausible operations, which are far from being fully understood (although the stronger argument, supported by Watumull (p.c.), is that this abstraction is in fact very close to the reality). Instead, we deal with an abstraction of the architecture, the Turing machine, and build our theory of economy on top of this architecture, using time and space complexity in the form of Big-O notation. As a result, our phonology can contain any computational operation (since it is a Turing machine), *and* we have a means of evaluating which (constraints on) operations are more plausible, based on time and space complexity.

So, let's put this into practice. For simplicity, I primarily compare (classic) OT and the rule-based phonology espoused by Samuels (2009), which draws much from Hale & Reiss (2008) and also from Minimalist syntax (Chomsky, 1993 *et seq.*). In line with the descriptive-explanatory, epistemological-ontological tension detailed in Section 2, there are two important domains of each theory: everyday (adult) use of phonological computations, and corresponding mechanisms to acquire this competence.

Samuels (2009) proposes three arguably domain-general procedures needed for phonology in everyday use: SEARCH, which provides a means for a probe to establish a relation with its goal; COPY, which copies a feature from a goal to a probe; and DELETE, which removes an element from the derivation. Note that there is no ADD, which would violate the Inclusiveness Principle (Chomsky, 2001) and greatly add to the complexity of the derivation, since constraints on what can and cannot be added – and what exactly is the source of the additions – would be required[10]. In any case, COPY and

---

[10] Chomsky (1995, p. 238) overtly claims that Inclusiveness is "radically false" at the sensorimotor interface. That this may not be necessary, as per Samuels' (2009) model, is therefore an interesting counter-hypothesis. Distributed Morphology (Halle & Marantz, 1993) introduces further complications here, but these lie beyond the scope of this article.

Received: 30/07/2020
Revisions: 10/02/2022
Accepted: 06/07/2023

DELETE are both $O(1)$, whereas SEARCH is worst-case $O(n)$, since the algorithm may need to check every element sequentially to see if it can serve as a goal. Every phonological algorithm can use only these operations albeit in any permutation, so the analysis of the worst-case time complexity of a derivation in this theory needs to find the slowest possible combination of these operations. Significantly, combinations of COPY and DELETE do not actually contribute to complexity on this metric, since $O(1+1+1) = O(1+1) = O(1)$, i.e., constant time, following the simplification rule dictated above. SEARCH is more problematic – one can conceive of a particularly troublesome derivation where each element in the derivation serves as a probe and needs to use SEARCH to find its goal. The worst-case scenario here is *n* searches, each of a progressively larger set as each element gets added, up to a maximum of n. In other words, this is mathematically a series, so its time complexity can be written as $O(\sum_{r=1}^{n} r)$, which expands to $O\left(\frac{1}{2}n(n+1)\right) = O(n^2)$. The worst-case time complexity of a derivation where every element can be a probe is therefore $O(n^2)$, which is very poor. This analysis makes it clear that placing a limit on n, such as with phase theory (Chomsky, 2000 *et seq.*; Samuels, 2009) and other theories of phonological locality (cf. Rose & Walker, 2011), is a computational imperative, lest the temporal burden of the derivation spiral out of control in quadratic time. Along these lines, the number of probes could also be limited. Indeed, this is the approach taken in Chomsky's (2008) model, where only syntactic 'phase heads' can serve as probes, entailing that each phase has only one probe. If such an analysis could be adapted to a Samuels-like phonological model, then the complexity of derivation is reduced to $O(n)$, equivalent to a single instance of SEARCH.

Moving onto OT, it must be mentioned that the computational complexity of OT has been much debated. For context, the use of asymptotic complexity in some form to analyse the complexity of theories of natural language dates back at least to Barton et al. (1987). It is also important to note that this approach does not come without criticism, as anticipated by Barton et al. (1987, section 1.4.1): 'Aren't the complexity results irrelevant because they apply to problems with arbitrarily long sentences and arbitrarily large dictionaries, while natural languages all deal with finite-sized problems?' Following the Turing programme for linguistics, as set out in Section 3, this is not a concern: the working, minimalist assumption is that linguistic competence *is* infinite, limited only by performance constraints. Idsardi (2006), adapting an earlier proof from Eisner (1997), presents a proof that the computational model provided by Classic OT is 'NP-hard', meaning that it is not computable in polynomial time on the size of the grammar. Kornai (2009) disputes the validity of the proof – to him, it has a 'fundamental flaw: [using] a mathematical method, asymptotic analysis, that relies in an essential fashion on the ability to create arbitrarily large problem instances' (Kornai, 2009, p. 702). Heinz et al. (2009) also dispute Idsardi's (2006) result, claiming that finite constraints on grammars eliminate the intractability problem. This, again, appears to hint at OT's concern with 'possible' languages as opposed to 'humanly computable' languages, as discussed in Section 2. To borrow Hale & Reiss's (2008) thought experiment: if only the performance systems of humans were to change suddenly, allowing computations of much greater time and space complexity to take place, would the computational procedure for human language be any different? Of course, by the very nature of the competence/performance distinction, it would not. This entails that the capacity for infinity be built into the grammar itself, a proposal that must be taken seriously. Indeed, similar arguments are presented by Roberts et al. (in press), and by Watumull (2015) for syntax. Returning to the OT issue, Heinz et al.'s (2009) arguments lose some of their weight, because they rely on a finitude that is incompatible with the Turing approach to phonological computation. Taking seriously the phonological enterprise as defined by Hale & Reiss (2008) makes such an approach especially appealing.

Even leaving these concerns aside, the OT algorithm intuitively appears slow compared to that of Samuels (2009). Since GEN is not considered part of the computation in OT, it will be assumed as $O(1)$, although this is of course a problematic omission from the theory, as discussed in Eisner's (1997) formalism. Each output of GEN then undergoes a constraint evaluation under EVAL. For m outputs and n constraints this algorithm takes $O(mn)$ in the worst-case scenario. Unlike in Samuels' theory, however, each instance of EVAL, for each output of GEN, can be calculated in parallel, as they are independent from one another until each has reached its end. One could argue that, with parallel operations, the time complexity could be reduced to $O(n)$. This kind of optimisation doesn't create a level playing field, however: we could easily apply parallel optimisation to Samuels' algorithm too, which might be able to reduce the time complexity. OT theorists may point to this as a drawback of the Turing machine approach – it only allows one operation to occur at a time. As Watumull (2012) explicitly points out, however, there is nothing inherently against there being *multiple* Turing machine instances in the mind, which interact in a kind of network. A Turing machine is, after all, simply a model for computation – it is clearly possible to have multiple computations taking place simultaneously. This remains a fundamentally different proposal from connectionism, which assumes instead that the nodes in the network are simple processing devices (Buckner & Garson, 2019).

Additionally, Big-O notation actually hides some of the nuances. The OT algorithm requires, without fail, $O(mn)$ consistently – effectively quadratic, which is very inefficient. The order of magnitude of operations is also vastly different: a very low estimate for the number of OT constraints needed (n) is 30, and GEN presumably needs about that many outputs, if not more, in order to cover every representational possibility. A more realistic estimate for the size of CON is given by Ashely et al. (2010) in their large-scale review of the OT literature, in which they conclude that there are at least 1,666 constraints – making this computation unreasonable, even leaving the acquisition algorithm aside. On the other hand, the locality constraints built into Samuels' RBP, in the form of phases, mean that n is probably never that large at all.

In light of these facts, OT is clearly more computationally complex than Samuels' brand of SFP. Assuming equivalent empirical coverage (descriptive adequacy), the demands of genuine explanation dictate that the computationally simpler theory prevail. This is not, of course, an innocent assumption, but a sufficient empirical review is beyond the scope of this article.

Traditional RBP, using *SPE*-style rewrite rules, isn't necessarily much better, despite at least being computable in polynomial time. Indeed, despite the appearance of being derivational, it is actually fundamentally different from Samuels' system, the latter being based on features and operations that build the derivation stepwise. Indeed, in some sense, Samuels' operations are more like Stampe's processes: they apply iteratively, sequentially, and locally. Traditional cyclic rules are linear, but necessarily run externally to the derivation, rather than from within. Each rule needs to search through the working representation in order to find its context (left-hand side). This ends up being just as computationally complex as OT in the fully developed adult system, merely replacing 'constraints' with 'rules'. Using Samuels' system appears much simpler from this perspective: the derivation only searches through (a limited portion of) the working representation for a single feature when triggered by SEARCH, instead of searching for a particular, potentially highly complex configuration for each rule in each cycle.

The facts become more interesting when looking at language acquisition, where one can argue that time and space complexity play an even greater role, because of the performance constraints imposed on the child. The goal of the child is to reduce the hypothesis space – to one, ideally. Problematically, it is a fact of our reality that there are always an infinite number of hypotheses

Received: 30/07/2020
Revisions: 10/02/2022
Accepted: 06/07/2023

compatible with any dataset. The Halting Problem inevitably arises here: how does the child (in particular the phonological Turing machine) 'know' when it is finished, or even if it will finish? Following Roberts & Watumull (2015), in the tradition extending from Chomsky, in turn extending from Leibniz: 'it cannot be proved that an "unsuccessful" search of the hypothesis space of possible explanations has been exhaustive' (p. 220). Even more crucially, this applies to *both* the Linguist's problem and the Human's problem – in the same way that the Child needs UG in order to evaluate their theory of *a* language, the Linguist needs some way of limiting their possible theories of language. The logical conclusion for the child is the presence of a large epistemologically prior toolkit to deal with the input in order to establish the most likely hypothesis (Chomsky, 1986), and this is likely through some kind of Bayesian inductive reasoning (Tenenbaum & Griffiths, 2001) or some system compatible with Yang's Tolerance Principle (Yang, 2016) – indeed, this is where the necessity of UG as defined in Section 2 above becomes evident. The Linguist, however, does not apparently have access to UG in order to solve the Linguist's problem. One typical Chomskyan answer to this is presented by Ludlow (2011): the Linguist *does* have access to UG, in the form of intuition. The other forms of evidence presented are also relevant, as ways of 'seeing into' competence. The more radical suggestion advocated here is that consideration of computational complexity should be added to the list of sources of evidence. In the spirit of the biolinguistic programme, it seems relatively inoffensive to attribute this tentatively as a consequence of evolution. Time is a reasonable evolutionarily grounded pressure – fast communication saves lives. This also applies in the acquisition domain: rapid acquisition of language allows the organism to focus on other important things, like survival and procreation – having the LAD atrophy immediately could be beneficial evolutionarily. This also links to third factors like attention: since there is a practically unbounded amount of sensory information theoretically available to a human at any single moment, filters on this input are essential for learning anything. Consequently, the sooner that the detailed attention to language needed to learn it is taken out of the picture, the sooner other preoccupations can take hold. Thinking requires time and energy, both of which are naturally in limited supply. There is thus every chance that cognitive economy is a highly weighted constraint on the Bayesian algorithm (or such) that inductively hypothesises and evaluates a theory of a language. As Chomsky (2004) hypothesises, if the crucial aspect of language is Merge, then it is highly probable that one can go *beyond* explanatory adequacy in suggesting that at least this component of language *is* optimal, in line with the Strong Minimalist Thesis. Following this line of reasoning, it is not at all unreasonable, as done in Section 3, to propose the Turing machine abstraction, and thus the computational phonological component on top of this, especially seeing as researchers from otherwise in places diametrically opposed persuasions agree on phonology being necessarily computational (Berent, 2013; cf. Hale & Reiss, 2008; Samuels, 2009)[11].

In following, it becomes necessary to look at what computational operations are required of the child in each learning algorithm. On the one hand, OT champions so-called 'factorial typology': the number of possible languages is the number of permutations of the constraint set CON, which, assuming strict ordering only, is $n!$[12]. With the earlier estimate for n as 30, the child is thus faced with an inconceivably huge innately endowed possibility set. With every new piece of input, CON is reordered accordingly, through an algorithm like that of Smolensky (1996). Reordering constraints in this way is

---

[11] A reviewer notes that there are of course interesting alternatives to this 'mind-as-computer' perspective. One example is that of *embodied cognition* (Shapiro, 2011), in which a full explanation of faculties of the mind/brain like language may necessarily need to refer to physical properties of humans, blurring the line between competence and performance and making the ontological nature of mathematical abstractions less clear. A full review of such alternatives is beyond the scope of this article, but see Reiss (2009) and Vukovic et al. (2017) for examples from phonology.

[12] With Stochastic OT (Boersma & Hayes, 2001), the picture is greatly complexified, as the strict ordering requirement is abandoned.

a relatively complex process: indeed, one could argue that it may even take factorial $O(n!)$ time complexity, as each constraint needs to be compared to all of the others in turn, every time a reordering takes place – this would therefore be completely intractable, although it is possible that this algorithm can be optimised. Idsardi's (2006) proof directly concerns the ranking problem – assuming that CON is not set in stone, which it evidently cannot be for acquisition to take place, he proves that this ranking problem is indeed 'NP-hard', meaning that the solution does not belong to the class of algorithms computable in polynomial time. In other words, the ranking problem is computationally intractable. Furthermore, Hale & Reiss (2008) show that the proposed learning algorithm is empirically problematic. As an aside, it is likely that extensions to Classic OT, such as Stochastic OT (Boersma & Hayes, 2001) and the theoretical propositions of Jarosz (2019), heighten the computational requirements, as a consequence of the need to calculate weighted constraints, at least when computations are considered within the context of the phonological Turing machine (and not, for instance, a connectionist or neural network model, cf. Gallistel & King, 2010, for general comments on computation and connectionism).

I propose that there are at least two further tentative ways out here. Firstly, eliminate the innateness requirement of CON – constraints are then acquired along with their ranking. This radical step cannot be taken lightly and comes with its own whole host of problems. One is that it is difficult to constrain a theory of emergent constraints, since wellformedness constraints are inherently defined negatively – as Reiss (2008) posits, what's to stop a child positing the constraint NOBANANA? Whilst this particular example is evidently *reductio ad absurdum*, it is undeniable that an infinite set of constraints can be proposed to account for any input data set – e.g., **#hta* is valid for English but surely inconceivable as a part of the grammar. This option to save OT thus eliminates any potential explanatory power, by virtue of the nature of language acquisition. It is also not clear whether this solves the complexity issue – clearly, the space complexity is reduced, as 'unused' constraints will never need to be postulated, but the generation of constraints appears very similar to the NP-hard problem proposed by Idsardi (2006). The second way I can conceive of saving OT, which is at first appearance more modest, is to constrain the learning algorithm to make more conservative adjustments to CON at a time, which do not involve factorial comparison. It is difficult to see, however, how factorial typology holds up, if fundamental alterations such as these are required in the name of computational economy. These facts appear especially problematic when contrasted with parametric systems which instead assume a set of binary parameters within UG, which would give a maximum of 2^n possible languages. This is still large, but an exponential typology is far easier to deal with than OT's factorial one. Furthermore, optimisations can be made to an exponential typology, following for instance Elan Dresher's Contrastive Hierarchy (Dresher, 2015) and Ian Roberts' Parameter Hierarchy (Roberts & Holmberg, 2010; Roberts, 2019), by assembling parameter hierarchies in the form of binary branching trees defined by implicational relations. Such a decision tree can be easily implemented as a conditionally branching Turing program, i.e. 'If C go to X, else go to Y', where X and Y are positions on the tape. Indeed, a particular implementation of the type of parameter hierarchy advanced by Roberts & Holmberg (2010) is provided by Bazalgette (2015) for certain syntactic parameters.

In any case, both approaches require some definition of the 'features' which populate the constraints or the parameter hierarchy. Much substance-free work (*pace* Reiss, 2017) allows for the existence of so-called emergent features, which allow features to be generated as necessary by the child to account for the input; an example implementation of this is presented by Mielke (2008). Following this line, a template system for emergent features is incredibly useful: it is a given that a strict template drastically reduces a hypothesis space, with exceedingly minimal overhead since the template itself is

Received: 30/07/2020
Revisions: 10/02/2022
Accepted: 06/07/2023

typically fairly simple. This is thus very much in line with Biberauer (2019), who for syntactic features proposes the template [*uF*, *iF*], where *uF* is a feature uninterpretable at the interfaces and *iF* one that is interpretable (see Chomsky, 1995, p. 277 for a technical description of 'interpretability'). Phonological features could have their own template, or, perhaps more optimally, they could even recycle the syntactic template, or the syntax could recycle a phonological template. The latter suggestion plays into a crucial discussion on the ontology of language: are the most important resources available to the child and the clues they look for syntactic, phonological, or conceptual-intentional? There is almost certainly, and perhaps necessarily, evidence for all three, and so perhaps the feature template is indeed something much more fundamental. Instead of interpretability being defined at the sensorimotor and conceptual-intentional interfaces used by syntax, the relevant interface for phonology is, naturally, the phonetics – or, more precisely, the device which *transduces* the phonology to a form legible by the phonetic module. It is entirely conceivable that some features are uninterpretable in this sense since phonology under SFP is arbitrary and symbolic. Nevertheless, this is likely an empirical matter to decide. The 'recycling' option will prove especially appealing if it appears to capture templatic effects across domains, which would appear to suggest that the broad concept of 'features' does have some kind of fundamental basis as a unified phenomenon. This would be particularly important for the Turing machine approach, since, as discussed for Samuels' SFP above, features are often the trigger for computations, and it would also reveal something more widely about the computational capacity of the mind if this theory is on the right track.

5 Conclusions

A theory of phonology – indeed any theory of I-language – must be epistemologically adequate, in that it must *describe* the universal human knowledge of language, and individual human knowledge of *a* language. It must also be ontologically adequate, being able to *explain* how this knowledge is acquired by the child, through some combination of genetics, emergence, and inductive Bayesian learning. It must predict the *phenomena* involved in language, which may be obscured by the data and the theories imposed thereupon. The prediction of the theory presented from Section 3 onwards is that the Turing machine is a sufficient and arguably necessary abstraction of the computational procedure for phonology, and for language, and indeed for the mind generally. The evidence needed to test this prediction will come from considerations of child language acquisition, for instance as presented in Section 2, as well as from evolutionary considerations, as are popular within the biolinguistic programme. Following Chomsky (1993 *et seq.*) there is increasing evidence that the language faculty is in some sense optimal. Accepting this conclusion has significant implications for the evaluation of theories of this faculty. Section 3 advocates a strongly derivational approach, and the suggestion in Section 4 is that theories should consider the implementation of the primitive operations and data structures they propose in the Turing machine. Computational complexity would be the evaluation metric between different options. Samuels' (2009) theory is taken as a case study in phonology, as it makes use of clearly defined computational procedures. It would be rash to suggest that this is the most optimal theory, however, and this is in part due to ambiguity about the precise nature of 'features'. The apparent potential of this approach and the corresponding philosophical deficit in OT directly contradicts McCarthy's (1988) proclamation that rules follow from representations; on the contrary, the clearer the picture of the operations and derivational procedures used by the mind, the clearer the picture of the actual representations they build will become.

6 References

Altenberg, E. P., & Vago, R. M. (1983). Theoretical implications of an error analysis of second language phonology production. *Language Learning*, 33(4), 427-447.

Anderson, S. R. (1981). Why phonology isn't 'natural'. *Linguistic Inquiry*, 12(4), 493-539.

Ashley, K. C., Disch, L., Ford, D. C., MacSaveny, E., Parker, S., Unseth, C., Williams, A. M., Wong, R., Yoder, B. (2010). How many constraints are there? A preliminary inventory of OT phonological constraints. *Occasional Papers in Applied Linguistics*, 9, 1-36.

Bach, E., & Harms, R. T. (1972). How do languages get crazy rules. In R. P. Stockwell & R. K. S. Macaulay (Eds.), *Linguistic change and generative theory* (pp. 1-21). Bloomington: Indiana University Press.

Barton, G. E., Berwick, R. C., & Ristad, E. S. (1987). *Computational complexity and natural language*. Cambridge, MA: MIT Press.

Bazalgette, T. (2015). *Algorithmic acquisition of focus parameters* (PhD Dissertation). University of Cambridge. Retrieved 23/07/2020 from *lingbuzz*: <https://ling.auf.net/lingbuzz/003006>

Berent, I. (2013). *The phonological mind*. Cambridge, UK: Cambridge University Press. Cambridge Core.

Berko, J. (1958). The child's learning of English morphology. *Word*, 14(2-3), 150-177.

Biberauer, T. (2019). Factors 2 and 3: Towards a principled approach. *Catalan Journal of Linguistics*, (special issue), 45-88.

Bley-Vroman, R. (2009). The evolving context of the fundamental difference hypothesis. *Studies in Second Language Acquisition*, 31(2), 175-198.

Boersma, P., & Hayes, B. P. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, 32(1), 45-86.

Bogen, J., & Woodward, J. (1988). Saving the phenomena. *The Philosophical Review*, 97(3), 303-352.

Brody, M. (2002). On the status of representations and derivations. In S. D. Epstein & T. D. Seely (Eds.), *Derivation and Explanation in the Minimalist Program* (pp. 19-41). Oxford, UK: Blackwell.

Buckner, C., & Garson, J. (Autumn 2019). Connectionism. (E. N. Zalta, Ed.) Retrieved 21/07/2020 from *The Stanford Encyclopedia of Philosophy*: <https://plato.stanford.edu/archives/fall2019/entries/connectionism/>

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1964). *Current issues in linguistic theory*. The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1966). *Cartesian linguistics: A chapter in the history of rationalist thought*. New York, NY: Harper and Row.

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

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York, NY: Praeger.

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

Chomsky, N. (1997). Questions: São Paulo. *DELTA: Documentação e Estudos Em Linguística Teórica e Aplicada*, 13, 123-128.

Chomsky, N. (2001). Derivation by phase. In M. J. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 1-52). Cambridge, MA: MIT Press.

Received: 30/07/2020
Revisions: 10/02/2022
Accepted: 06/07/2023

Chomsky, N. (2004). Beyond explanatory adequacy. In A. Belletti (Ed.), *Structures and beyond* (pp. 104-131). Oxford, UK: Oxford University Press.

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36(1), 1-22.

Chomsky, N. (2006). *Language and mind: Third edition*. Cambridge, UK: Cambridge University Press.

Chomsky, N. (2007). Approaching UG from below. In U. Sauerland & H.-M. Gärtner (Eds.), *Interfaces + Recursion = Language?* (pp. 1-30). Berlin: De Gruyter Mouton.

Chomsky, N. (2008). On phases. In R. Freidin, C. P. Otero, & M. L. Zubizarrete (Eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*. Cambridge, MA: MIT Press.

Chomsky, N. (2014). Minimal recursion: Exploring the prospects. In T. Roeper & M. Speas (Eds.), *Recursion: Complexity in cognition* (pp. 1-15). Cham, Switzerland: Springer.

Chomsky, N. (2021). Minimalism: Where are we now, and where can we hope to go? *Gengo Kenkyu*, 160, 1-41.

Chomsky, N., Gallego, Á. J., & Ott, D. (2019). Generative grammar and the faculty of language: Insights, questions, and challenges. *Catalan Journal of Linguistics*, 229-261.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. Cambridge, MA: MIT Press.

Dresher, B. E. (2015). The motivation for contrastive feature hierarchies in phonology. *Linguistic Variation*, 15(1), 1-40.

Eisner, J. (1997). Efficient generation in primitive Optimality Theory. *35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics*, 313-320. Madrid, Spain: Association for Computational Linguistics.

Eisner, J. (2000). Easy and hard constraint ranking in OT: Algorithms and complexity. *Proceedings of the Fifth Workshop of the ACL Special Interest Group in Computational Phonology*, 22-33. Centre Universitaire, Luxembourg: International Committee on Computational Linguistics.

Epstein, S. D., & Seely, T. D. (2002). Introduction: On the quest for explanation. In S. D. Epstein & T. D. Seely (Eds.), *Derivation and explanation in the Minimalist Program* (pp. 1-18). Oxford, UK: Blackwell.

Ferguson, C. A. (1986). Discovering sound units and constructing sound systems: It’s child’s play. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 36-57). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gallistel, C. R., & King, A. P. (2010). *Memory and the computational brain: why cognitive science will transform neuroscience*. Malden, MA: Wiley-Blackwell.

Gervain, J., & Mehler, J. (2010). Speech perception and language acquisition in the first year of life. *Annual Review of Psychology*, 61(1), 191-218.

Gombert, J. E. (1997). Metalinguistic development in first-language acquisition. In L. Van Lier & D. Corson (Eds.), *Encyclopedia of language and education: Knowledge about language* (pp. 43-51). Dordrecht: Springer.

Hale, M., & Reiss, C. (1998). Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry*, 29(4), 656-683.

Hale, M., & Reiss, C. (2000). “Substance abuse” and “dysfunctionalism”: Current trends in phonology. *Linguistic Inquiry*, 31(1), 157-169.

Hale, M., & Reiss, C. (2008). *The phonological enterprise*. Oxford, UK: Oxford University Press.

Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), *The view from Building 20* (pp. 111-176). Cambridge, MA: MIT Press.

Heinz, J., Kobele, G. M., & Riggle, J. (2009). Evaluating the complexity of Optimality Theory. *Linguistic Inquiry*, 40(2), 277-288.

Huang, S. (2020, January 16). What is big O notation explained: Space and time complexity. Retrieved 09/02/2022, from *freeCodeCamp*: <https://www.freecodecamp.org/news/big-o-notation-why-it-matters-and-why-it-doesnt-1674cfa8a23c/>

Hume, E. (2011). Markedness. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology* (pp. 79-106). Malden, MA: Wiley-Blackwell.

Idsardi, W. J. (2006). A simple proof that Optimality Theory Is computationally intractable. *Linguistic Inquiry*, 37(2), 271-275.

Jarosz, G. (2019). Computational modeling of phonological learning. *Annual Review of Linguistics*, 5(1), 67-90

Kager, R. (1999). *Optimality Theory*. Cambridge, UK: Cambridge University Press.

Kelly, T. (Winter 2016). Evidence. (E. N. Zalta, Ed.) Retrieved 28/10/2019 from *The Stanford Encyclopedia of Philosophy*: <https://plato.stanford.edu/archives/win2016/entries/evidence/>

Kenstowicz, M. J. (1994). *Phonology in generative grammar*. Oxford, UK: Blackwell.

Kenstowicz, M. J., & Kisseberth, C. (1979). *Generative phonology*. New York, NY: Academic Press.

Knuth, D. E. (1997). *The art of computer programming* (3rd ed.). Boston, MA: Addison-Wesley.

Kornai, A. (2009). The complexity of phonology. *Linguistic Inquiry*, 40(4), 701-712.

Lewontin, R. C. (2000). *The triple helix: Gene, organism, and environment*. Cambridge, MA: Harvard University Press.

Liberman, M., & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff & R. T. Oehrle (Eds.), *Language sound structure: Studies in phonology presented to Morris Halle by his teacher and students* (pp. 157-233). Cambridge, MA: MIT Press.

Marantz, A. (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics*, 4(2), 14.

McCarthy, J. J. (1988). Feature geometry and dependency: A review. *Phonetica*, 45(2-4), 84-108.

McCarthy, J. J., & Prince, A. (1993). *Prosodic Morphology: Constraint interaction and satisfaction* (Technical Report No. RuCCS-TR-3). Retrieved from *Rutgers University Center for Cognitive Science*: <http://roa.rutgers.edu/files/482-1201/482-1201-MCCARTHY-0-1.PDF>

Mielke, J. (2008). *The emergence of distinctive features*. Oxford, UK: Oxford University Press.

Miller, B., Myler, N., & Vaux, B. (2016). Phonology in Universal Grammar. In I. Roberts (Ed.), *The Oxford Handbook of Universal Grammar* (pp. 152-182). Oxford, UK: Oxford University Press.

Ohala, J. J. (1986). Consumer's guide to evidence in phonology. *Phonology*, 3, 3-26.

Ohala, J. J. (2005). Phonetic explanations for sound patterns: Implications for grammars of competence. In W. J. Hardcastle & J. M. Beck (Eds.), *A figure of speech: A festschrift for John Laver*. New York, NY: Routledge.

Prince, A., & Smolensky, P. (1993). *Optimality Theory: Constraint interaction in generative grammar* (Technical Report No. RuCCS-TR-2/CU-CS-696-93). Retrieved 10/02/2022 from *Rutgers University Center for Cognitive Science*: http://roa.rutgers.edu/files/537-0802/537-0802-PRINCE-0-0.PDF

Pylyshyn, Z. W. (1984). *Computation and cognition: Toward a foundation for cognitive science*. Cambridge, MA: MIT Press.

Received: 30/07/2020
Revisions: 10/02/2022
Accepted: 06/07/2023

Reiss, C. (2008). Constraining the learning path without constraints, or the OCP and NoBanana. In B. Vaux & A. Nevins (Eds.), *Rules, constraints, and phonological phenomena*. Oxford, UK: Oxford University Press.

Reiss, C. (2009). Intermodular explanation in cognitive science: An example from phonology. In D. Dedrick & L. Trick (Eds.), *Computation, cognition, and Pylyshyn* (pp. 201-217). Cambridge, MA: MIT Press.

Reiss, C. (2017). Substance free phonology. In S. J. Hannahs & A. Bosch (Eds.), *The Routledge Handbook of Phonological Theory* (pp. 425-452). London: Routledge.

Rizzi, L. (2009). Movement and concepts of locality. In M. Piattelli-Palmarini, P. Salaburu, & J. Uriagereka (Eds.), *Of minds and language: A dialogue with Noam Chomsky in the Basque Country* (pp. 155-168). Oxford, UK: Oxford University Press.

Roberts, I. (2019). *Parameter hierarchies and Universal Grammar*. Oxford, UK: Oxford University Press.

Roberts, I., Watumull, J., & Chomsky, N. (in press). Universal Grammar. In D. Vakoch (Ed.), *Xenolinguistics: Toward a science of extraterrestrial language*. Oxford, UK: Oxford University Press.

Rose, S., & Walker, R. (2011). Harmony systems. In J. A. Goldsmith, J. Riggle, & A. C. L. Yu (Eds.), *The handbook of phonological theory* (2nd ed., pp. 240-290). Oxford, UK: Blackwell.

Samuels, B. D. (2009). *The structure of phonological theory* (PhD Dissertation). Harvard University, Cambridge, MA. Retrieved 12/04/2020 from *lingbuzz*: https://ling.auf.net/lingbuzz/000863

Shapiro, L. (2011). *Embodied cognition*. London: Routledge.

Smith, N. V. (1973). *The acquisition of phonology: A case study*. Cambridge, UK: Cambridge University Press.

Smith, N. V. (2009). *Acquiring phonology: A cross-generational case-study*. Cambridge, UK: Cambridge University Press.

Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry*, 27(4), 720-731.

Stampe, D. (1979). *A dissertation on natural phonology*. New York, NY: Garland.

Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24(4), 629-640.

Thornton, S. (Winter 2019). Karl Popper. (E. N. Zalta Ed.) Retrieved 15/07/2020 from *The Stanford Encyclopedia of Philosophy*: https://plato.stanford.edu/archives/win2019/entries/popper/

Trubetzkoy, N. (1958 [1969]). *Principles of phonology* (C. A. M. Baltaxe, Trans.). Berkeley, CA: University of California Press.

Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society, 2nd series* 42(1), 230-265.

Vaux, B., & Watumull, J. (2012). *The phonological Turing machine*. University of Cambridge / MIT. Unpublished manuscript.

Vukovic, N., Feurra, M., Shpektor, A., Myachykov, A., & Shtyrov, Y. (2017). Primary motor cortex functionally contributes to language comprehension: An online rTMS study. *Neuropsychologia*, 96, 222-229.

Watumull, J. (2012). A Turing program for linguistic theory. *Biolinguistics*, 6(2), 222-245.

Watumull, J. (2015). *The linguistic Turing Machine* (PhD Dissertation). University of Cambridge.

Weizenbaum, J. (1983). ELIZA: a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 26(1), 23-28.

Yang, C. (2016). *The price of linguistic productivity: How children learn to break the rules of language*. Cambridge, MA: MIT Press.

About the author

Louis Van Steene is a former MPhil Linguistics (by thesis) student at the University of Cambridge. His research spanned the fields of formal syntax and phonology, in particular regarding issues in the realm of computation and formalisation.

Contact: louis@lvs.me.uk

Acknowledgements

I would like to thank Bert Vaux, Jeffrey Watumull, Ian Roberts, and Bridget Samuels for helpful discussion and/or comments on previous versions of this paper. This article began as an undergraduate examination essay, and as such I would also like to thank all of my supervisors during my time at Cambridge for helping shape my thought (if perhaps often unintentionally) in the fascinating direction it has taken, with special thanks going to my former Director of Studies, Theresa Biberauer. Finally, I would like to thank two anonymous reviewers for their comments and suggestions, as well as the team at JoULAB for their meticulous guidance and support.

Funding Disclosure: Part of this work was funded by the Harding Distinguished Postgraduate Scholars Programme Leverage Scheme.

Received: 30/07/2020
Revisions: 10/02/2022
Accepted: 06/07/2023