FORMAL IS NATURAL: TOWARD AN ECOLOGICAL PHONOLOGY

Dafydd Gibbon, Universität Bielefeld (final submitted version, 2009-04-30)

Table of Contents

Abstract

Naturalism Phonology (NP) has a history of opposition to abstractness, to generative linguistics, to formalist approaches, and differs from these in its strong focus on external rather than distributional, structural evidential domains. But evidence domains are orthogonal to empirical and formal methods, and, like formalist theories such as Optimality Theory (OT), the pedigree of NP includes structuralist and generative phonology. In an analysis which is sympathetic to both NP and OT, this contribution examines the relation between NP and OT, analyses a classic OT case study of syllabification in Tashlhiyt Berber, and presents computational linguistic analyses of this case, as well as of English syllable phonotactics and of tone language tonotactics. The contribution advocates an opening towards these methods, and the adoption of explicit, consistent, precise, complete and sound formal criteria for theories, which enable an exact interpretation in terms of operational models and computational implementations, and practical applications. The general frame of reference is a the Ecological Cycle in theory formation, from clarification of the domain through theory construction, interpretation with a model, evaluation and application in the original evidential domain, with payback to the language community from which the evidence was gained.

1. Background and objectives: the natural and the formal[1]

1.1. The natural and the formal

Natural human behaviour, including natural phonetic behaviour, is to some extent regular, to some extent chaotic. To the extent that it is regular, natural behaviour can be described with precise categorial or statistical rules. These rules

---

[1] Some of the specific points and examples in this contribution have been presented and published elsewhere; the overall framework is new, however.

can be formulated in plain text form, with the disadvantage that the formulations may be more ambiguous, vague or general for the reader than they were intended to be by the writer. Or they can be formulated with systematic notations with intuitive interpretations, still with the disadvantage that these interpretations may be more ambiguous, vague or general for the speaker than they wereintended to be by the writer. Or they can be formulated using a notation which is – preferably unobtrusively – interpretable in terms of a well-defined algebra or logic, with the advantages that their interpretations are not only explicit but also consistent, precise, complete and sound, and can be checked with automatic support if necessary, saving much time, effort, and uncertainty.

Taking 'natural' to mean that a theory is empirically well-grounded, and 'formal' to mean that a theory is based on well-defined methods of reasoning, five criteria which relate the formal and the natural can be formulated, without which there can be no strict confirmation of refutation of a theory, or of a particular description based on this theory:

1. *Explicitness*: a representation with a formal logical or mathematical interpretation. Plain text, the top-level metalanguage in which *explicanda* and modelling conventions are formulated, is in general the least explicit (and therefore open to the most misunderstandings); a systematic but nevertheless ad hoc alphanumeric or graphic symbolism, the most common in descriptive linguistics, is less open to misunderstandings, and a well-defined formalism such as those used in fields from logical semantics to phonetic signal processing is the most explicit, and the least prone to misunderstanding.
2. *Consistency*: a representation without contradictions or paraphrases. Plain text may provide more entertaining reading, but leaves much to the idiosyncratic interpretation skills of the readers; an ad hoc symbolism also leaves much to individual interpretation, but for a formalism consistency is a necessary property.
3. *Precision*: a representation which corresponds in reproduceable detail to the empirical domain under investigation. The model in terms of which the statements of a theory are interpreted, has an exact and detailed relation to empirical reality.
4. *Completeness*: a representation which corresponds to all the empirical data under investigation. There are no counterexamples which the theory should but cannot predict.
5. *Soundness*: a representation which corresponds to only the empirical data under investigation. Junk entities and properties which are not actually found in the data are not predicted.

This set of criteria will be referred to in the following discussion as 'the ECPCS criteria', and a formal approach will be taken to be one which can be shown to fulfil this criterion set. A formal approach in this sense provides an intellectual toolset which is a *necessary* but by no means *sufficient* condition for a good theory.

Formal linguistics is not simply 'the description of the forms of languages' but 'the use of formal methods in linguistics' (Gibbon 2005). Functional varieties of linguistics, the description of language in contexts of use, are not exempt from the ECPCS criteria, and have often conformed to these criteria. The best known examples are from Halliday's Systemic Functional Grammar, the basis of his functional semiotics, which has been used since the 1970s in Artificial Intelligence software (starting with Winograd 1971), an impossibility without formalisation. The software used by linguists and phoneticians is another example: there is a highly formal underlying lexicon theory behind *Toolbox*, and highly formal signal processing and signal-symbol mapping theories behind and *Praat*, *Wavesurfer*, *Transcriber* and *ELAN*.

The same is true of the text linguistic theories which underlie the document models and stylesheets for text and hypertext description used on the internet and in word processors, and which were developed by archivists, librarians and by the linguistic research departments of major software providers.

The present contribution does actually concentrate on language forms, however, for a good reason. In a 'synchronic' temporal perspective, there is an epistemological priority of form over function: if there are functions, there are forms, but not necessarily vice versa, otherwise one would not be able to formulate nonsense words and sentences, or have the experience of not understanding what is being said. In a 'diachronic' temporal perspective, functions without forms are also conceivable, but in a teleological sense, as goals which are yet to be instrumentally fulfiled by forms.

The concept of a Finite State Model (FSM) plays a significant role in this study; the term is used as a cover term for more technical concepts such as Finite State Automaton, Finite State Transition Network, Finite State Transducer, and related concepts such as Regular Grammar, Type 3 Formal Grammar, Regular expression, which will not be explained in detail. The linguistically relevant property of FSMs is that they describe relationships between linear structures, rather than hierarchical tree structures or more complex graphs; FSMs have been used extensively in various computational linguistic subdisciplines, including computational morphology, phonology and lexicography, as well as in statistical language models in the human language technologies.

Besond the modelling of language functions and forms there is a more general theoretical framework which requires the ECPCS criteria to be respected: the Ecological Cycle of theory development, from a pretheoretical *explicandum* clarification and statement of modelling conventions through theory construction to interpretation in terms of a model of a fragment of reality, evaluation and application in ‘payback’ to the sources of the data used in theory development.

1.2. Strategy of presentation

Section 2 deals with ‘the beauty’ and ‘the beast’ and their encounters. In particular, the relation between Natural Phonology (NP) and Optimality Theory (OT) is discussed, and a seminal case from the early days of OT is dissected in some detail. The stance taken is sympathetic to *both* NP *and* OT, and these theories are regarded as overlapping in some respects, complementary in others, and not as competitors, except in the usual Kuhnian paradigm senses of competing for attention, money, gifted young scholars, and the creation of mutual citation cartells, etc., which are not germane to the present argument. The central point in this and the following sections is that both NP and OT use collections of operations on representations as ‘rules’, ‘processes’ and ‘constraints’, but do not demonstrate explicitly how these operations or constraints are connected in a coherent system – the Saussurean ‘un système où tout se tient’. The cases to be discussed are both concerned with compositional, syntagmatic constraints – phonotactics and tonotactics.

In Section 3, the connectedness criterion is applied to the issue of syllable representation, with English as an example. In view of the complexity of syllable phonotactics in English there is a need for a detailed representation as a basis for defining the domains for phonological rules. A detailed FSM for defining the placement of phonological rules is presented and discussed as a model for applying the ECPCS criteria.

Section 4 deals with a further issue, of system connectedness: tonotactics in tone languages. Two cases are discussed in respect of the value of this type of representation for prosodic typology: tone terracing in West African tone languages, and tone sandhi in Mandarin Chinese. It is shown that differences in prosodic typology are reflected in the topologies of the respective FSMs which describe these patterns.

Finally, in Section 4, the conclusions, and a perspective on future mutual benefits of cooperation between the orthogonal dimensions of naturalness and formalism in Natural Phonology and other approaches is focussed within the Ecological Cycle of *explicandum* clarification, theory, model, empirical evaluation, assessment in application, and revision.

2. The beauty and the beast

2.1. The natural beauty meets the formal beast

Natural is beautiful, formal is the beast. The aim of this section is to stimulate discussion and modernisation of the concept of naturalism as a universal principle in post-generative linguistics: requirements for theory evaluation are extended beyond traditional notions of internal and external evidence, and assigned instrumental roles in a utilitarian ecology of science. For expository purposes I will associate Natural Phonology with ‘the beauty’ and Optimality Theory (and other formal approaches) with ‘the beast’, without approbatory or derogatory intention in each case.

The perspective of ‘science for science’s sake’, perhaps represented most notably by the the epistemologies of positivism, logical empiricism and critical rationalism, is a popular one, but in many ways unrealistic. Even Kuhnian paradigm theory and more recent approaches to the philosophy of science are limited in this respect. In linguistics, the view is often coupled with a restriction to ‘internal evidence’, i.e. evidence from judgments on the structure of language and on the distribution of smaller language forms within larger language forms. In contrast, the empirical foundation of Natural Phonology (and Natural Linguistics in general) is distinguished by its emphasis on ‘external evidence’, i.e. by taking evidence from language use in context. However, the Ecological Cycle does not close: ‘products’ are rarely put back into these source environments in operational applications and evaluations of linguistic results.

A key factor in understanding the development of science, which has been determinedly ignored in many quarters, including linguistics but less so in phonetics, is technology. We do not need to look at physics, with its telescopes and large hadron colliders, to confirm this. The birth of modern phonetics in the second half of the 19th century, for example, can arguably be assigned to two applications oriented pushes. The first applications push was in technology, with the successful marketing of the telephone by the Scottish speech therapist, phonetician and engineer Alexander Graham Bell. This push was preempted by an even earlier, though perhaps less spectacular push, by Alexander Graham Bell’s father, Alexander Melville Bell who developed a phonetic transcription system, ‘Visible Speech’ as professor of elocution at University College, London, establishing a line of research there which was later continued by Sweet, Jones, Gimson and Wells. The second applications push was in foreign langage teaching, not unrelated to the research of the older Bell, with the founding of the International Phonetic Association and its journal *Le Maître Phonétique*, later the *Journal of the IPA*. And in modern phonetics the completion of the Ecological Cycle from theory to application and back tends to be is the rule rather than the exception.

Likewise, many recent insights in phonetics during the past two decades owe a great deal to the technology push which first made it possible to used computers to analyse very large quantities of speech data and create precise models of speech articulation, and then took computers out of laboratories and put them on every phonetician's desk as PCs and laptops. Part of this technology push were personal and community initiatives which created sophisticated free software for users of these computers. The combined effect has been a creative cycle of scientific ideas which go back into the technology and from there push intellectual developments forward in an Ecological Cycle. Technology is not causal in this development: the cause lies in the inventiveness of the human mind. Technology is instrumental and utilitarian, a necessary, though not a sufficient condition for science, from the use of tools for text production to the use of tools for data discovery, analysis, evaluation and application. Computation is likewise a necessary, but not a sufficient condition for technology. And, in the humanities, neither technologies nor computational models need be close mirrors of the workings of the mind.

This claim is intended to provoke discussion by taking up the central feature of external evidence, which is closely associated with Natural Phonology. Theory-internal exploitation of external evidence is not sufficient, and in any case theories have full explanatory status only if they fulfil the ECPCS criteria of explicitness, consistency, precision, completeness and soundness. This is not to say that only a partial conformity to the Ecological Cycle is not a valid goal, nor is it to claim that only technological applications demonstrate the value of science: it is a powerful heuristic role which technology plays in the Ecological Cycle, not a theoretical role.

2.2. Optimal beasts: preferences, optimality, defaults and overrides

A relatively recent innovation in modern linguistics is the introduction of the 'other things being equal' principle, under various names: 'elsewhere principle', 'typicality', 'default inference', 'mutatis mutandis' condition. Under this principle, some generic strategy, event, object, property is prioritised unless other more specific considerations countermand this priority. A central theoretical concept in NP is the 'preference': other things being equal, a certain rule of fortition or lenition may be preferred, i.e. may apply, otherwise this may not be the case. In a similar vein, other things being equal, in OT, a more general constraint may be violated if necessary. An example of default-override relations is English /r/, typically realised as a voiced sound. However, this is just the 'default', 'elsewhere' or otherwise' case, which is overridden by a more specific condition: in on syllable onsets, where /r/ is typically realised as unvoiced after a homorganic voiceless stop, as in [tr̥iː]. Other things being equal, English plurals have an /s/ suffix, unless attached to a stem with a final voiced sound or a sibilant; other things being equal, English sentences have SVO word order, unless the sentence is interrogative, etc.

Both OT ranked constraints and NP preferences can in principle be formalised using a default logic, in which a particular general, typical, default case may be overridden by a particular, more specialised case. Indeed, in the case of Optimality Theory the tableau method of deciding on optimal representations originates in a standard proof method in constraint logic.

This is apparently where the similarities between NP and OT end, not because the one approach is wrong and the other is right, but because they are orthogonal in respect of several *tertia comparationis*. A seletion of these dimensions of comparison is listed in Table 1.

In summary: OT corresponds most closely to the ECPCS criteria, but NP is better placed within the Ecological Cycle. The key issue is that Natural Phonology is a descriptive theory within Natural Linguistics, with an explanatory framework which allows for 'external evidence', but in which preference selection is not clearly formalised. Optimality Theory, on the other hand, is a very precise evaluation theory based on formalised proof methods which is designed to make it easy to falsify a given constraint configuration, but the empirical basis for the constraints and their optimal ordering are unclear. These are fundamental epistemological differences.

Both theories also share (with each other and with other theories of phonology in the post-generative tradition) the problem of the unclear epistemological status of the 'input' to their rule and constraint systems.

The 'input' in standard OT is an abstract, underlying lexical representation (as it is in NP), and it is not clear how to transfer this coherently to external domains such as the 'input' to a child in language acquisition, as has been suggested in psycholinguistic work (Barlow 2001), or to other external domains such as pathological deterioration or recovery in aphasia, or to a stage in the history of the language. The term 'input' referring to the lexicon in relation to a phonology is at a very far remove indeed from the term 'input' referring to the signal source in a communication channel. The formal constraint resolution method of OT can, of course, be used in any domain[2], but in all fairness, this is not OT, it is constraint logic.

Table 1: Some *tertia comparationis* for Natural Phonology and Optimality Theory.

tertia comparationis	*Natural Phonology*	*Optimality Theory*
Logic:	default (intuitive, e.g. rules are ordered and preferential and precede processes)	default constraint logic (e.g. markedness constraints outrank faithfulness constraints)
Inference:	unconstrained ordered derivational rules (rewriting or substitution rules), also analogical argumentation (cf. 'figure-ground' metaphor)	modus ponens (all inputs of type *A* are blocked by constraint *p*; input *x* is an *A*; therefore *x* is blocked by constraint *p*), implemented as a tableau-based constraint resolution procedure from formal logic
Generalisations:	rules before processes	faithfulness before markedness constraints
Representations:	plain text + symbolisation with unconstrained derivational rule notation	formalisation as ordered constraints defaults
Explanation	descriptive + final (teleological)	hypothetico-deductive
Empirical evidence	internal (distributional) + external (diachronic, pathological, ...)	internal (with recent applications to external domains)

The OT formal mechanism corresponds to a well-known search strategy of 'generate and test' (or in this case 'generate and filter'). The negative (markedness) and positive (faithfulness) constraints can in principle be compared with rules and processes, and perhaps also fortition and lenition preferences respectively, including the ordering principles of NP (rules before processes) and OT (markedness before faithfulness); these dimensions of comparison do not match completely, however.

Finally, a very simple test. How many syllables are there in English? OT is in principle capable of predicting the size of this set: it is equal to the size of the output of the generator component, minus the size of the subsets filtered out by each constraint. However, NP is incapable of making this prediction, and it is unclear even how to start. Is this a trivial issue which is not worthy of the attention of a serious phonologist? Not at all, at least if a processing theory of external evidence is aimed at: it has a direct bearing on processing issues such as memory size and processing time in a theory of production and perception, and on the issue of learnability. The topic will be taken up in Section 3.

3. A hint for the beast: how to approach the beauty

3.1. A syllable segmentation problem: OT and a Finite State approach

As a starting point, an OT description is compared with a description in classic computational linguistic style, using a Finite State Model. In fact, OT can easily be regarded as a strategy for gently teaching computational methods in phonology; in this it has been highly successful, and there are now many computational tools to assist with OT analyses.

The first widely circulated OT account of a problem in phonotactics was in a discussion by Prince and Smolensky (1993, reprinted in McCarthy 2003) of the syllabification problem for Tashlhiyt Berber, which permits consonantal syllable nuclei. Although this account is old and OT has moved on much further, it is a classic and straightforward example of OT and still worth discussing. The discussion will be structured in terms of the Ecological Cycle: clarification of the explicandum, definition of modelling conventions for a fragment of reality, theory construction, derivation of hypotheses, comparison with the model, operational application, and revision as necessary.

3.2. Clarification of the task, modelling conventions, data set

The Tashlhiyt syllabification task is to segment lexical words into syllables, which is generally done in order to define contexts for phonological and prosodic rules. The general linguistic requirements for the modelling conventions are:

---

[2] I am grateful to Forugh Shustaryzadeh for discussions on this point.

1. a preference for syllables to have onsets (but the simpler the better) except (in the case of Tashlhiyt) phrase initially,
2. a preference for a syllable nucleus to have greater sonority than syllable margins,
3. a default/elsewhere concept for catching more specific contexts for rule or constraint application first, and more general default contexts later.

Modelling conventions involving preference requirements of this kind have been most clearly formulated in the Natural Phonology paradigm (e.g. Vennemann 1988). The rule or constraint ordering principle of specific-first-default-last (the first regular expression has the more specific nucleus specification) has been discussed in Generative Phonology as the 'elsewhere condition' and respectfully attributed to the Sanskrit grammarian Pāṇini (about 400 BC). However, as already noted, it is a general principle of default logic and plays a role in standard reasoning (cf. *if p and q then p,* though not the inverse), and is presumably as old as awareness of the human reasoning faculty.

The OT modelling convention for these preferences is to generate a set of strings with all possible syllable boundary placements between segments (the GEN component), and then to filter out the relevant ones using ordered (ranked) constraints (the EVAL component). One constraint prefers a syllable to have an onset (except phrase-initially), one which prefers a nucleus to have a higher sonority (e.g. a vowel rather than a consonant). Prince and Smolensky categorise the problem as a parsing problem; the analysis strategy for syllabification is a constraint resolution problem.

Table 2: Tabulation of McCarthy & Prince Tashlhiyt syllabification data (syllabification of the output set is indicated by dots).

Set of input strings	*Set of segmented output strings*	*Search space size*
ratkti	.ra.tk.ti.	$2^5 = 32$
bddl	.bd.dl.	$2^3 = 8$
maratgt	.ma.ra.tgt.	$2^6 = 64$
tftkt	.tf.tkt.	$2^4 = 16$
txznt	.tx.zmt.	$2^4 = 16$
tzmt	.tzmt.	$2^3 = 8$
tmzh	.tmzh.	$2^3 = 8$
trglt	.tr.glt.	$2^4 = 16$
ildi	.il.di.	$2^3 = 8$
ratlult	.rat.lult.	$2^6 = 64$
trba	.tr.ba.	$2^3 = 8$
txznas	.txz.nas.	$2^5 = 32$

The modelling conventions from a computational linguistic perspective using FSMs are a little different. The categorisation 'parsing' is actually loose terminology: syllabification is a segmentation task, prior to parsing, known as 'tokenisation' or 'lexical analysis'. Parsing is in general the assignment of a tree structure labelled with categories, so, strictly speaking, no parsing is involved.

The segmentation problem is better categorised as a search problem: the search space (set of all possible candidates) of all possible segmentations can be traversed with any of a number of possible search strategies. For an input string of length *n*, the size of the search space is 2^{n-1} because there is a binary choice between each segment whether to place a syllable boundary or not. For a string such as *ratlult* of length 7, the search space is therefore $2^{7-1} = 2^6 = 64$ possible segmentations (initial and final boundaries are obligatory, so do not affect the search problem). For illustration, a shorter example is taken: *bddl* with length 4, and search space size of $24^{-1} = 2^3 = 8$:

{ b.d.d.l, b.d.dl, b.dd.l, b.ddl , bd.d.l, bd.dl, bdd.l, bddl }

In Table 2, the entire data set of inputs and outputs referred to by Prince and Smolensky is listed, together with the size of the search space for each item.[3] The set is tiny; too tiny, in fact, to prove the validity of the theory, and no doubt only intended for expository purposes.

The modelling convention used in OT is not completely explicit (the GEN component is inadequately characterised). Seen as a search strategy, known as a 'generate and test' procedure, the simplest possible search strategy. It is not a particularly 'intelligent' strategy (in a technical sense), in that all items in the search space have to be investigated, and for some problems with large search spaces it is not feasible to generate all members of the search

---

[3] Prince and Smolensky do not present a solution for the pair *txznakkw*, *.txz.nakk*^{w}*.>*, so it is not included. This case is simple to solve with an additional coda constraint.

space. Indeed, for searching in infinite search spaces (such as the set of sentences for natural languages) it is impossible to generate the whole search space all at once, so the generate and test strategy fails immediately for this kind of task. The search procedure used in the CLFS approach is described below.

However, the search task involved in syllabification involves *deterministic* search, that is, search with only one outcome, as opposed to *nondeterministic* search, that is, search with more than one outcome. This being the case, the OT machinery is overkill for this simple task, which, as already noted, was no doubt selected mainly for expository reasons rather than as a matter of principle.

3.3. Components of the OT theory, and hypothesis derivation

The components of the OT theory, as alread noted, are the GEN component which generates the entire search space, and the EVAL component which reduces the search space to the correct candidate (deterministic search) or candidates (nondeterministic search). An OT theory can be represented from a formal point of view as two functions, GEN and EVAL applied in sequence:

$$\text{GEN}(\textit{lexicon}) \rightarrow \textit{searchspace}$$

$$\text{EVAL}(\textit{searchspace, constraintset}) \rightarrow \textit{output}$$

or as a composed function of the GEN and EVAL components:

$$\text{EVAL}(\text{GEN}(\textit{lexicon}), \textit{constraintset}) \rightarrow \textit{output}$$

Table 3: Partial constraint resolution tableau for selected search space element. Columns: first, violation counter result; second, search space of candidates; others, constraint violation marks (Prince & Smolensky 2003:15). [Left column with GEN function and the top row of elucidations added. DG]

<table>
<tr><td rowspan="7"><i>GEN(lexicon) →</i></td><td><i>violation counter</i></td><td><i>search space</i></td><td colspan="3"><i>EVAL (a constraint grammar)</i></td></tr>
<tr><td colspan="2">Candidates</td><td>Ons</td><td>Hnuc</td><td>Comments</td></tr>
<tr><td>☞</td><td>.tX.zNt.</td><td></td><td>n x</td><td>optimal</td></tr>
<tr><td></td><td>.Tx.zNt.</td><td></td><td>n t !</td><td>|n| = |n|, |t| < |x|</td></tr>
<tr><td></td><td>.tXz.nT.</td><td></td><td>x ! t</td><td>|x| < |n|, t irrelevant</td></tr>
<tr><td></td><td>.txZ.Nt.</td><td>*!</td><td>z n</td><td>Hnuc irrelevant</td></tr>
<tr><td></td><td>.X.Z.N.T.</td><td>*!****</td><td>n z x t t</td><td>Hnuc irrelevant</td></tr>
</table>

Once the GEN component has generated the full search space, in the EVAL component, the constraints (here, onset preference Ons and higher sonority preference Hnuc) are applied in a predetermined order (Ons before Hnuc), to each element of the search space, and if the constraint is violated (i.e. the element is not valid) then a violation mark is assigned. This is known as a 'breadth first' search procedure; Prince and Smolensky call it 'parallel', though strictly speaking this is not quite accurate. The list of items in the search space and the set of constraints to be applied to them is represented as a table, called a *tableau*, the term for a common constraint resolution configuration in constraint logic. The number of cells in the tableau is therefore the size of the search space times the number of constraints, i.e. for an input of length n, and m constraints $2^{n-1}m$. Finally, the element with the smallest number of violation marks is selected.[4] A tableau with accepted and rejected syllabifications is illustrated in Table 3. In the tableau, an asterisk '*' signifies a constraint violation, '!' signifies whether the violation is absolute and not just a preference, shaded cells mean that the constraint in that column is inapplicable. The optimal (accepted) item in the search space is indicated by the pointing finger '☞', and in this case does not violate either of the two constraints, while other candidates produce violations.

The OT hypothesis formation process takes place in several derivational steps: one for the GEN component, one for each element of the search space, in each case multiplied by the number of constraints, and one per element of the search space for the violation counter. And, yes, the search procedure is, formally speaking, a derivation (i.e. a sequence of steps in a reasoning process), and like other derivational models has intermediate results, despite ubiquitous polemics

[4] The counting algorithm which is somewhat laboriously introduced by Prince and Smolensky is actually a simple list length comparison algorithm normally found in elementary introductions to programming.

in OT against ‘derivational approaches’, and it is still a derivation even if there is in effect only one rule of inference, *modus ponens*.

Note that the OT description contains a number of selective aspects, which mean that the correctness of the analysis is not actually demonstrated: first, because of the tiny data set; second, because for each input item a only a small, arbitrary subset of the search space is intuitively selected; third, the exact processing procedure is only characterised fragmentarily (this will be taken up in the section on the FSM approach, below).

Evidently, the Prince and Smolensky OT machinery is massive overkill in regard to this simple segmentation problem. For Tashlhiyt syllabification, the CLFS theory only needs a small part of the sonority hierarchy, concerning the consonant-vowel obstruent-consonant relationships (in the latter case, also permitting other obstruents as possible nuclei.

3.4. Components of the FSM theory, and hypothesis derivation

The components of the FSM theory are a little different. A traditional representation of a grammar is as a function:

grammar(*lexicon*) → *output*

But this is not enough for a the syllabification task (or indeed any actual task), which in the present case requires a *tokenise* operation as an additional operation over well as a *lexicon* and *categories*:

tokenise(*lexicon, grammar*) → *output*

In this description, the *grammar* is a rule set which corresponds in principle to the constraint set of OT. In the tokeniser, the search space is explicitly structured in three kinds of choice:

1. by the order of processing the input strings (left-right, right-left, chunk spotting, etc.),
2. by the choice of top-down (e.g. starting with the category of syllable and comparing its variants with the input) vs. a bottom-up search procedure (i.e. comparing substrings of the input with syllable variants until the syllable structure is found),
3. by the choice of breadth-first search schedule (explore all alternatives according to one criterion, rule or constraint before proceeding to the next criterion, rule or constraint).

It does not usually matter too much which of these choices is made, particularly if the search is deterministic. The end result will be the same (with a few restrictions in more complex cases). OT remains silent on these issues, except to describe the procedure as ‘parallel’; actually ‘breadth-first’ is meant, and not ‘parallel’ in the sense of ‘temporally simultaneous’. Still, a choice has to be made. The FSM theory described here has left-right, bottom-up, breadth-first processing. In other words, the input strings are processed from left to right, boundary criterion matching is done directly on the input string, and each criterion (rule, constraint) are applied to each string before proceding to the next criterion.

The *grammar* component has just two rules which express the same linguistic requirements on modelling conventions which the OT constraints express: prefer syllables with onsets (except phrase-initially), and prefer nuclei which are more sonorant then their neighbours. Both rules express the obligatoriness of onsets (initial no-onset syllables are left inexplicit), and the sonority ordering of consonant and nucleus; the preference for higher sonority nuclei is expressed by the ordering of rules with higher sonority nucleus first:

1. Any consonant-vowel sequence is analysed as *consonant^nucleus* sequence.
2. Any obstruent-consonant sequence is analysed as *consonant^nucleus* sequence.
3. The beginning of a word is the beginning of a syllable.
4. The end of a word is the beginning of a syllable.

And that is all. The specification can be reformulated in terms which describe the OT notation:

1. Set a dot in front of any consonant-vowel sequence.
2. Set a dot in front of any obstruent-consonant sequence.
3. Set a dot at the beginning of the word if there is not one there already.
4. Set a dot at the end of the word.

A declarative, OT-like constraint-based formulation would be:

1. Scatter dots between each character, and at the start and end.
2. Accept sequences of *dot ^ consonant ^ vowel*.
3. Accept remaining sequences of *dot ^ obstruent ^ consonant*.
4. Reject the rest.

Table 4: Unix shell script implementation for Tashlhiyt Berber syllabification FSM.

```
           Unix shell script                          Inline comments
#!/bin/bash
VOW="[aiu]"                           # Vowel phoneme set
CON="[ptkbdgfshvzmnrl]"               # Consonant phoneme set
OBS="[ptkbdgfshvz]"                   # Obstruent phoneme set
NODOT="[^.]"                          # Set of anything which is not a dot
LEFTEDGE="^"                          # Left word boundar
RIGHTEDGE="$"                         # Right word boundary
ADDLEFTDOT=".&"                       # Insert dot left of sequence
ADDRIGHTDOT="&."                      # Insert dot right of sequence
sed "s/$CON$VOW/$ADDLEFTDOT/g         # Insert dot before cons vow sequence
     s/$OBS$CON/$ADDLEFTDOT/g         # Insert dot before obstruent consonant sequence
     s/$RIGHTEDGE/$ADDRIGHTDOT/g      # Insert dot at end
     s/$LEFTEDGE$NODOT/$ADDLEFTDOT/g” # If no dot at beginning, insert
```

All these definitions produce equivalent results and fulfil the ECPCS criteria: *explicit* not only with regard to rules, but also with regard to the processing schedule; *consistent*, in being contradiction-free and paraphrase-free; *precise*, in describing the data in detail; *complete*, in covering all the data, and *sound* in not producing junk descriptions. Fulfilment of the ECPCS criteria permits definition of an operational model and a computational implementation (cf. what may be the simplest possible implementation in Table 4).But a caveat is necessary: the FSM model is intended to capture just the data presented by Prince and Smolensky, which is a different task from trying to capture the entire Tashlhiyt syllable inventory. This follows the spirit of their presentation, however, but in neither case can completeness or soundness be taken as shown for the entire lexicon of Tashlhiyt.

3.5. Evaluation

The OT approach can, a little whimsically, be characterised as a gentle introduction to computational methods in phonology. The approach does not attain the standard of fully meeting the ECPCS criteria, though, and there are a number of other unresolved issues apart from this:

1. The procedural specifications are inexplicit.
2. The data set is too small.
3. No criteria for the selection of the search space subset are given.
4. No distinction is made between the linguistic requirements for modelling conventions, which are already very well known in Natural Linguistics, on the one hand, and on the other hand the actual modelling conventions of OT.
5. Criteria for a full operational and computer-implementable model of the OT tableau method is not discussed, and applications areas are not envisaged. Implementation of such a model would not be very difficult (similar formalisms and implementations have been around for a long time), except for the fact that the procedural specifications are inexplicit.
6. The description of the elementary list length comparison algorithm for checking numbers of violations is rather laborious, indeed irrelevant in view of the lack of fully explicit specifications for the rest of the method.
7. The machinery proposed for Tashlhiyt Berber is overkill: simpler machineries are possible.

But in defence of OT it must be pointed out that though the data set discussed is tiny, the descriptive mechanism is not actually intended to be limited to this data set, but to apply fully to Tashlhiyt, and, beyond this, to apply – with the aid of the more general formulations – to other problems than syllabification, and other languages than Tashlhiyt. These claims about universality, generalisability and typological usefulness cannot be tested just using such a small amount of data, however.

Finally, by far the most interesting point to be made, however, has been formulated by Karttunen (1998): OT descriptions can be translated straightforwardly into a cascade of FSMs (plus a default operation), i.e. a series of FSMs whose output provides the input for the next, so strategies for operationalisation and implementation are not far away. qually interesting: Kay & Kaplan (1994) had previously shown that FSM cascades of the types discussed in the present contribution can be composed into a single FSM, a finding which also applies to the EVAL component of OT phonology, but also possibly to the GEN component if, as seems likely, it can be expressed as an FSM.

4. Beast approaches beauty: modelling phonotactics

4.1. À la recherche d'un système où tout se tient

It is a characteristic of both OT and NP that, like other linguistic descriptions of the generative type, they formulate sets of rules whose interdependencies are not fully explicit. Partial dependencies between, such as the cyclical rule application principles, feeding and bleeding output/input relations, the 'elsewhere' principle, the strata of Lexical Phonology, the ordering of markedness and faithfulness constraints in OT, the ordering of rules and processes in NP, are steps along the way. However, in order to show in detail how the rules are related, a connected representation is required, a phonotactic grammar of syllables and phonological words. In general, connections between rules are left implicit or, like the notion of morpheme structure rule, fragmentary. Exceptions in this respect are Vennemann's study of preferences in syllable phonotactics (1988), and, in principle, Dziubalska-Kołaczyk's rhythm-theoretic approach to phonotactics (2002), though both stop short of a full integration into a formalised phonological theory.

With this background, the aim of the present section is to demonstrate the feasibility of a complete, connected formal phonotactic system, based on the simple formalism of Finite State Automata, which provides explicit contexts for attaching phonological rules of the allophonic, contextual conditioning type.

4.2. A selection of phonotactic models

Syllable structure is taken as the point of departure here but without wishing to enter into the debate in Dziubalskan NP about whether the notion of syllable has a firm ontological or epistemological basis. This is an issue for linguistic typology; there are at least clear cases, and it is these which are meant here.

Syllable structure is conventionally represented in various ways, which will be dealt with here from a generic point of view, rather than as specific cases from specific publications.

1. One common representation lists constraints on segment class sequences: *V*, *CV*, *CVC*, *VC*, *CCCVCCC*, ...
2. Another defines constraints on segment properties in sequence, e.g. for English #*x*[stop][liquid], *x* = /s/
3. A third defines syllable structure in terms of sonority differences: $\text{son}_{\text{initial\_edge}} < \text{son}_{\text{peak}} > \text{son}_{\text{final\_edge}}$

A fourth prefers nested structures, represented as trees or bracketings, such as (Onset (Nucleus Coda)).

In more sophisticated versions, the notations are reasonably well-suited to describing selected properties of syllable structure. But they have in common that they are incomplete, each in a different way:

1. The *CVC* notation type lacks generalisations over constraints between specific sound subclasses.
2. The segment property constraint set, comparable with traditional phonological rules, provides only a fragment of an overall structure as context.
3. The sonority relation type, like the first, requires supplementation with generalisations over specific sound types and more detailed positions in the syllable.
4. The tree structure type requires supplementing with linear constraints which cross the branching structure.

The notations are *locally* explicit and precise. But they leave open the issues of consistency, completeness and soundness, which require a more *global* form of explicitness within the holistic context of an overall working system. So how, precisely, do we know that the descriptions are consistent, complete and sound, particularly when it comes to creating collections of specific rules and constraints?

An appropriate answer to these questions is to have a formalisation which will enable inconsistencies (e.g. due to rule and constraint ordering) to be identified on a principled and exhaustive basis, by combining each separate constraint into a single, connected coherent single system with holistic as well as atomistic properties. This is the accepted procedure in morphology and syntax: a formal grammar is designed with the task of enumerating all and only words, phrases, sentences, etc. of a language, i.e. explicitly addressing the formal criteria of completeness and soundness. Although 'phonology is different' in respect of its ordering constraints (Bromberger & Halle 1989), nevertheless it is subject to the same formalisation criteria as the rest of language, and thus a phonotactic grammar is needed in addition.

In the previous section, a simple test was formulated, to which no known phonological theory has an answer: How many syllables are there in English? Although linguists tend to fight shy of numbers, and apparently odd questions like this are not taken as serious linguistic issues, they are of general importance for a theory of phonological processing which takes memory access constraints and production and perception procedures into account.

The first step in ansering this question (and many other, even more interesting questions besides), is to write a syllable grammar. This task initially seems to be straightforward and uninteresting: the set of syllables is finite because syllables have a clear, small finite limit on their length (unlike words and sentences). Syllables could, therefore, simply be listed.

But the issue is not trivial, first because no generalisations would be expressed, second because syllable length may range from 2 in *CV* languages to about 8 in English or German which are C^nVC^m languages (depending on the details of the analysis of diphthongs and affricates). A CV language with, say, 15 consonants and 10 vowels, has maximally is 15 • 10=150, and there may be constraints on combining *C*s and *V*s, further limiting this number.

In a *CVC* language, assuming 15 consonants and 10 vowels, and no restrictions on consonant occurrence in onset and coda, the size of the set of potential syllables is already 2250. However, in practice there will be strong restrictions, particularly on coda consonants.

In a C^nVC^m language, the situation is clearly more complex. On the one hand, the increased number of syntagmatic positions raises the number of regular potential syllables exponentially. An exponential curve is not maintained, however, because typically there are strong restrictions on consonant occurrence. Based on detailed formalisations such as those of Berndsen (1998) and Belz (2000), the number of potential monomorphemic syllables in English is around 25,000, with a number of borderline cases. The number of actual syllables attested in the English lexicon is usually estimated at less than 10,000.

Limiting investigations to the set to actual syllables precludes explanation of how new syllables such as '(to) gyre', '(a/to) fax', '(to) skype', and many brand names, are invented; generalising over actual syllables shows how this is done: existing regular sub-sequences (onsets, nuclei and codas) are re-combined in new ways.

4.3. Generalising phonotactics

Syllable structure needs to be captured in a coherent, connected way which captures generalisations and thereby predicts new syllables.

One way to formulate a syllable grammar is to use a phrase structure grammar, as in many approaches to syntax and morphology. This kind of grammar is generally used for two reasons. First, it can define arbitrary kinds of recursion as well as constituent groups; however, for a finite language, recursion is not needed. Second, this kind of grammar has the advantage of showing syntagmatic clusters: groups which share mutual constraints, such as Noun Phrases in syntax, or Onsets, Rhymes, Nuclei and Codas in phonology. However, the tree structures generated by this kind of grammar have to be supplemented be a collection of constraints on linear cooccurrence, comparable to the selection restrictions and cooccurrence restrictions in syntax and morphology. Tree structures provide expressive generalisations over syntagmatic relations and paradigmatic classes, but syllable tree structures have finite depth, and can therefore be reduced to FSMs. There is certainly recursion of syllable structures in syllable sequences, but there is no centre-embedding recursion, as in subject relative clauses in English, but only rather simple iteration, i.e. unidirectional right or left recursion, and it is known that an FSM (or equivalent formalisms such as a *linear* or *regular* grammar[5]) is quite adequate.[6]

This argument demonstrates for both syllable structure and for syllable sequences that the appropriate connected and formally homogeneous representation for syllable structure is the FSM (and its formal equivalents). The other notations which have been discussed are very suitable for generalising over specific aspects of syllable structure, but do not show the entire topology of the syllable. A representation of a phonotactic grammar for uninflected English syllables is shown in Figure 1.

---

[5] An FSM represents a regular grammar, the most constrained grammar in the Chomsky hierarchy of formal languages, known as a Type 3 grammar. A phrase structure grammar is Type 2, a context-sensitive grammar (in the formal sense of the term, not in the way it is used in mainstream phonology) is Type 1, and an unrestricted grammar such as a transformational grammar is Type 0. Regular grammars describe the same languages as finite state automata (also known as finite state machines), finite state transition networks (cf. Figure 1) and regular expressions; these languages are more restricted than those described by phrase structure grammars (Gibbon 2005).

[6] An FSM is actually also sufficient for describing simple sentences (sentences with one finite verb and no conjunctions) in syntax: there is a maximum length and also a finite (though very large) number of such sentences. Why this point has not been made in the previous linguistic literature is unclear. Maybe it is too obvious? It is also true in morphology, as Koskenniemi (1983) has shown.

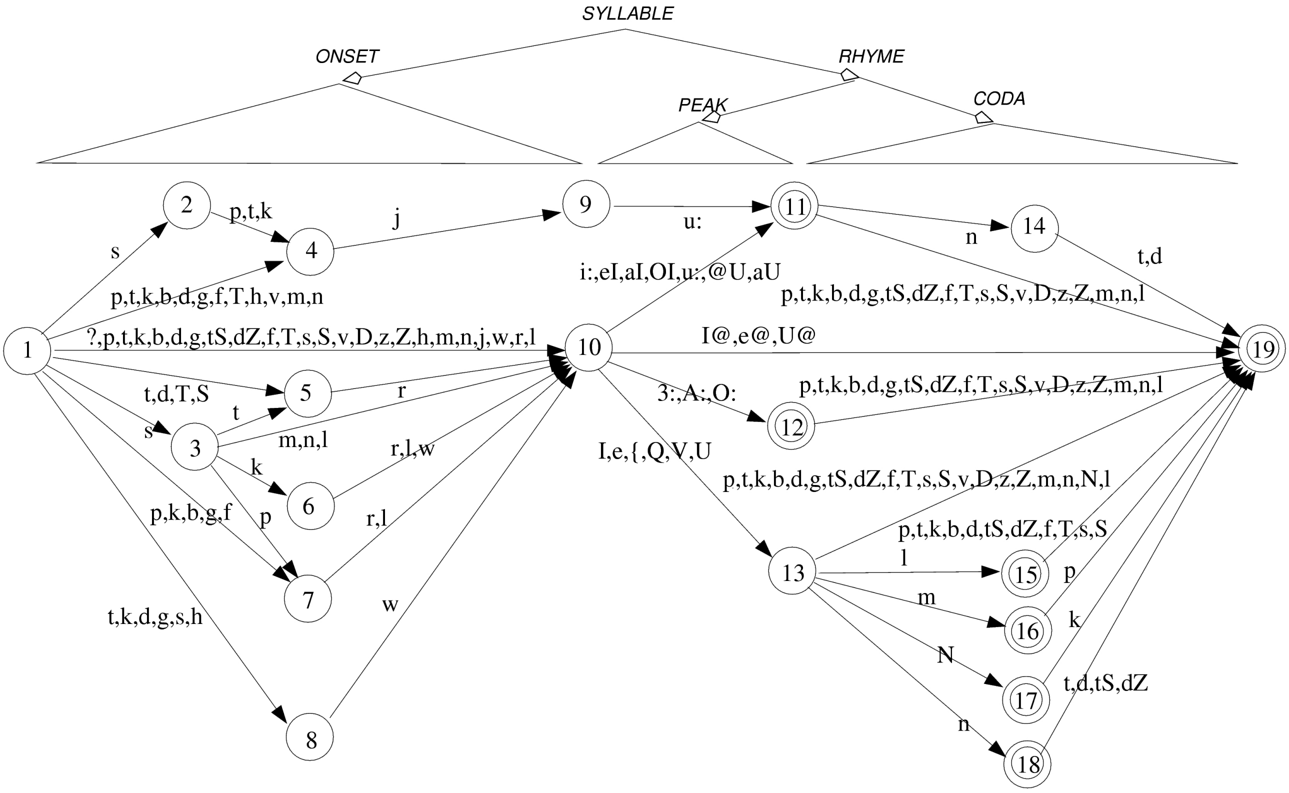

Figure 1: Finite state transition network for English syllable structure, with a conventional tree representation illustrating the main areas in the topology of the network. Each edge is labelled with set of symbols to avoid reproducing one transition per symbol; the symbols are rendered for simplicity in the keyboard and computation friendly SAMPA ASCII encoding of the International Phonetic Alphabet (Gibbon et al. 1997).

The FSM representation as a network (there are other standard visualisations) also has several advantages which are not immediately obvious, but which follow from the ECPCS formalisation criteria:

1. *Explicitness*:
    1. The mapping to tree-shaped Syllable, Onset, Rhyme, Nucleus, Coda notations is straightforward, each of these categories corresponding to an area in the network topology.The network shows the topology of the syllable structure of English directly.
    2. The onset structures and substructures (mutual constraints) are clear.
    3. The lack of constraints between onsets and rhymes is also clear: these two structures meet at a single node; the one exception is also clear: /u:/ after certain consonants in many English dialects.
    4. The complex constraints between Nucleus and Coda are clearly distinguished. The anomalous pattern with /u:/ after certain consonants is also clear.
    5. The subnetworks provide the contexts for defining dissimilation (in the onset) and assimilation (in the coda) between neighbouring segments.
2. *Consistency*:
    1. The notation is well-defined and unambiguous.
    2. The notation contains no paraphrases.
3. *Precision*:
    1. The linear left-to-right structure of the FSM appropriately represents the linear, non-hierarchical contexts required for identifying allophonic variation.
    2. The network can be split, if required, into the individual constraints of which it is constituted, to conform to the conventional expectations of phonologists.

    3. Non-local constraints are required, such as */s$C_i VC_i$/ (as in */spip/, */skak/, though cf. /stat/) where C_i denotes consonant occurences with identical properties, then these can be added most simply by duplicating subsections of the grammar (other more elegant methods are available).
4. *Completeness*:
    1. The network generates exactly 23597 potential syllables. In this respect the network is a suitable explication for an already highly constrained 'GEN' component of an OT description, the transitions of which can be decorated with individual phonological rules or constraints.
    2. The contexts for allophones of a given phoneme are completely defined by the position of the phoneme in the network, and can be explicitly generated by extending the finite state transition network to a related formal model, the finite state transducer (it would og too far to discuss this here).
    3. Syllable sequences are straightforwardly generated by iterating over the entire network.
    4. The network is nevertheless easily falsifiable on inspection (a positive feature) on the grounds of incompleteness, though these gaps are easily rectifiable:
        1. there is no path describing sequences such as /spju:/, /stju:/, /skju:/ or some Greek loan word onsets,
        2. bimorphemic (e.g. inflected) monosyllables are not included,
        3. weak syllables (e.g. lexically unstressed syllables as in syllable two of *happy* or *butter*) are not included.
5. *Soundness*:
    1. No impossible syllables are defined.
    2. The formal inelegance of using tree-based grammars to describe ambisyllabicity (where branches rejoin) is an artefact in an FSM: an intervocalic consonant simply appears linearly in a segment sequence with an intervening inter-syllable node in the network. In other words, there is no coercion of consonants into one syllable or the other at this level.

It is an elementary programming task to represent this kind of network in such a way that it can be interpreted by a well-defined procedure which is implementable in any programming language. The interpreter can be used either to accept or reject hypotheses about specific potential syllables, or to generate a complete set of potential syllables for checking. Attaching phonological rules or constraints to the transitions yields a transducer with the same kind of output as that targeted by OT and NP.

From the perspective of the Ecological Cycle it may be noted that this kind of linear representation is standard technology in computational phonology, computational morphology, and in the human language technologies, though it may seem somewhat alien to those used to isolated rule and constraint collections without such a reference structure. FSMs are used in many kinds of applications, from speech synthesis and speech recognition to morphological analysers and the spell-checkers provided in well-known word processors, and there is a very productive new field of study referred to as Finite State Technologies. A widely used and linguist-friendly software tool for regular grammars (which also converts standard phonological rule notations automatically into finite state automata) is the Xerox Research `xfst` package[7] (Beesley et al. 2003). Detailed operational accounts of FS phonology, with applications in speech recognition and machine learning, are also given by Carson-Berndsen (1998) and Belz (2001). Karttunen (1998) has shown that (with the exception of the default operation) OT constraint ranks can be translated into an FSM. A worthwhile challenge would be too check how far this can also be done for NP using such tools; this is not an idle speculation, but an educated guess.

5. Formal modelling of prosodic generalisations

5.1. Description: Tone in Niger-Congo languages

Finite State Modelling has frequently been used in prosody description, and operational naturalness has been approximated, for instance, in intonation generation in speech synthesis systems. The intonation literature cannot be reviewed here; suffice it to say that it has long been known, from work by Fujisaki in the 1960s and later, the IPO group in the 1970s, Pierrehumbert in the 1980s, and work by the present author from the 1980s to the present, that intonation systems can be represented by FSMs with contour, accent and boundary tone patterns: from initial boundary tones through iterated accentual tones to final boundary tones.

---

[7] The `xfst` software package was used to implement the network described in Figure 1, enabling the set of output syllables to be counted exactly. The package is distributed with Beesley & Karttunen (2003).

It is less well-known that similar considerations apply to the pitch patterns of tone languages. Unlike standard practice in syntax, the influential study of right-branching metrical trees for tone terracing by Clements (1981) did not include a tone grammar to generalise over specific trees as one would have expected in areas like syntax. It is not an accident that a formal property of FS grammars is that they can be used to construct right or left branching tree structures but not centre-embedded or mixed right and left branching tree structures. This insight was used to develop a FS model of tone terracing for Niger-Congo languages (Gibbon 1987, 2001, 2004).

Ibibio (Niger-Congo, specifically Delta Cross, ISO 639-3: ibb), simplifying slightly and ignoring lexical downstep, has two lexical tones, H and L. Nouns have lexical toneme contrast, verbs have templatic patterns and inflectional tone, and grammatical and word formation contexts provide structural templates to which tones are assigned. Because of these complex factors, tone sequences are more or less arbitrary from a tonological perspective alone, and could be represented by a 1-node network and two tonal transitions, one for H and one for L. However, tonal terracing requires phonetic interpretations for H and L which depend on the preceding linear context: HH, HL, LH, LL. The following phonetic conditions hold, in which lower case ‘h’ and ‘l’ stand for the phonetic correlates of the lexical tones ‘H’ and ‘L’. Languages differ in details:

1. HH: $h_1 = h_2$ (in other languages: $h_1 < h_2$ (upsweep), $h_1 > h_2$ downdrift)
2. LH: $l < h$ (downstepped H; in other languages: l raising, $m < h$; h lowering, $l < m$)
3. HL: $h >> l$ (in other languages: $h > m$, followed by lower l)
4. LL: $l_1 = l_2$ (in other languages: $l_1 > l_2$

The main point is that the step down from H to L is larger than the step up from L to H (downstep), leading to the characteristic Niger-Congo stepwise downward ‘terracing’ effect of the pitch pattern.

These relations between the lexical and the phonetic levels can be expressed in a coherent, connected FSM which meets the ECPCS criteria (cf. Figure 2).

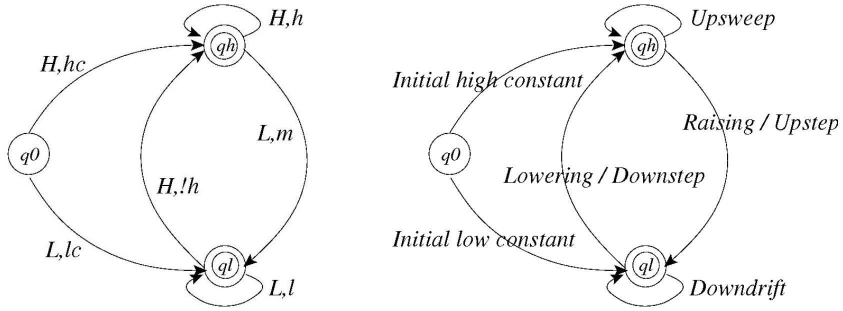

Figure 2: Oscillatory FSMs for terraced two tone languages: left, with phonetic tone interpretation (upper case = lexical tone, lower case = phonetic tone); right, with names of contextual conditioning types.

The FSM in the left hand part of Figure 2 has three nodes: a start node, from which H and L tones, phonetically interpreted with an approximately constant starting pitch, branch to a node which is the target for all H tones and a node which is the target for all L tones, respectively. Each transition in the network is decorated with a different ‘allotone’ rule. In the right hand part of Figure 2, traditional names from the tonology literature are provided on the transitions. A complete ‘terrace’ is represented formally by one transition cycle or oscillation from either a H node or a L node to the other node, then back to the same node. A sequence of like tones may be referred to as a ‘demi-terrace’.

The linguistic value of the FSM approach in the case of Niger-Congo terraced tone may be summarised in terms of the following typologically relevant properties:

1. a connected network with a complete set of contexts for tonological rules,
2. a clear framework for comparing languages typologically in terms of network structure (e.g. with more nodes to express more complex patterning),
3. a clear framework for interpreting tones in quantitative phonetic terms.

In traditional tonological descriptions, this level of generalisation is not attained.

5.2. Description: Tone in Mandarin

The question immediately arises of how far the tone FSM approach may be generalised to other language types. Jansche (1998) applied the approach to the Tianjin Mandarin (ISO 639-3: cmn) dialect by unifying tone sandhi rules into a connected graph (Figure 3) which describes potential tone sandhi sequences in this dialect. The unification procedure reveals marked typological differences between Niger-Congo and Mandarin tonal language types. More detailed comment on this is not called for in the present context.

The utility of this notation for linguistic typology becomes evident when the main properties of the FSMs for Niger-Congo tone terracing and Mandarin tone sequencing (tonal sandhi) are compared (Table 5).

Table 5: Comparison of structure of Niger-Congo and Mandarin FSMs.

Niger-Congo	*Mandarin*
Much smaller node & transition inventory.	Much larger node & transition inventory.
Clear major (terracing) loop or loops.	No major loop.
Local loop for sequences at each node.	Isolated loops.
Input-output relations are phonetic.	Input-output relations are morphophonemic.

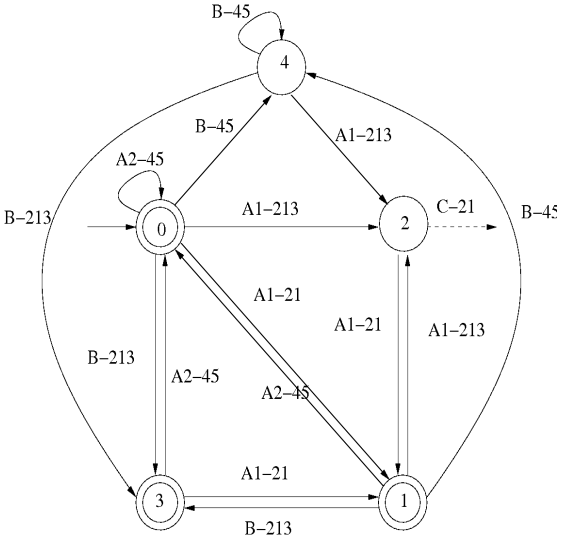

Figure 3: Tianjin Mandarin tone (after Jansche 1998).

5.3. The Ecological Cycle, application phase: speech synthesis of tone languages

An operational use of the FS model for empirical modelling of tonally annotated speech recordings, with mapping of phonetic tone categories to the quantitative operations of Pierrehumbert et al. (1984), is presented by Urua et al. (2003).

The Niger-Congo type of tonal FSA has been implemented in an operational speech synthesis prototype. Text-to-Speech (TTS) synthesis is one of the most useful HLT systems for use with non-literate communities, with typical applications in health, market and educational information, and fits well into the concept of the Ecological Cycle in the context of formal approaches, because the ECPCS criteria must be fulfilled in order for a speech synthesis system to be implemented. Current TTS development environments are theoretically, but generally not in practice, language independent. Voice creation is not easy, but adaptation of a voice in one language to another language or a different style or gender in the same language can provide an adequate platform for some purposes (Bachan et al. 2006).

Adaptation is plausible when languages are prosodically and phonemically similar but severe problems arise when languages are very dissimilar (e.g. ‘intonation languages’, for which TTS systems are typically developed, vs. ‘tone languages’, e.g. Ibibio). In the LLSTI project (Tucker et al. 2005), the adaptation procedure was applied to Ibibio. In the prototype for this project, the two-tone FS model was used (Gibbon et al. 2006). The generalisability of the approach was investigated by Govender et al. (2005).

Ongoing work also shows that the morphosyntactic functions of tone, for example in Ibibio, can also be modelled using FSMs (Gibbon et al. 2006). Initial studies show that these problems can also be generalised to other African tone languages, which have lexical (phonemic) tone but also morphemic and morphosyntactic tone, unlike most Asian tone languages (but cf. Hyman in press).

6. Summary, conclusion and outlook

An operationally extended notion of linguistic naturalism was introduced, based on the concept of an Ecological Cycle in theory formation, in which the tension between naturalistic and formalistic approaches was discussed in relation to selected design features of Natural Phonology and Optimality Theory. The Ecological Cycle has the phases clarification of an explicandum, theory formation, interpretation with a model, evaluation, application, and then restarting the cycle.

Further, a set of evaluation criteria for a good theory were introduced – explicitness, consistency, precision, completeness and soundness – were enumerated. Three areas of phonotactics (including tonotactics) were explored against this background: first, a classic case of an Optimality Theory analysis, with a Finite State Model as a simpler, more explicit alternative; second, English syllable structure; third, tone sequencing in tone languages.

A genuinely naturalistic approach to syllable or tonal structure (and of any language structure) is one which is able to follow the Ecological Cycle, and not shy away from using appropriate formalisation in the theory formation and interpretative model phases of the cycle, and appropriate technologies at the application phase of the cycle, in order to provide ‘payback’ to the context of language use and users from which the original external evidence came. Formal is natural. : explicit formalisation is not only a prerequisite for an empirically evaluable theory, but also a prerequisite for computation and, in turn, for technology and for the feedback which these can give to theory development, as well as providing further impetus for social development in the communities in which the technologies are deployed.

Within the present framework of the Ecological Cycle, the conflict between hard and fast solutions, which are fragile, and prioritised ‘on demand’ solutions, which degrade gracefully, is a key issue. The explicit use of preferences, overridable constraints and prioritisation is important in the practical contexts of everyday science: no-one wants to be trapped in an environment which is either safe or life-threatening and nothing in between – non-default escape routes are called for.

Discussions of scientific methodology in linguistics have often appealed to scientific revolutions. However, technological revolutions and their influence on science as facilitating instruments are just as important, both in conducting empirical and theoretical science, and in creating and disseminating the products of science, i.e. publications, quite apart from their effects on the general consumer domain.But we do not need to go back to Gutenberg and printing technology, or even to Archimedes, or Newtonian and Einsteinian mechanics, in order to illustrate the Ecological Cycle. Phonetics has perhaps been most significantly affected by the introduction of technological support for the validation of theoretical work, and in the introduction two ‘applicaitions pushes’ on phonetics were noted. Two examples from modern phonetic scenarios will illustrate the point further:

1. The Praat phonetics workbench (and many other software tools, but less so) has revolutionised the empirical study of acoustic phonetics and the creation of high quality phonetic data in large quantities by taking these activities out of specialised labs and on to the desks of thousands of phoneticians world-wide. Other software packages are either more specialised, or restricted to particular working environments. The relationship between what has been called ‘symbol phonetics’ and ‘signal phonetics’ has been (and is still being) further clarified by work with time-aligned annotation of transcription labels with speech signals to an extent which was not previously possible.
2. Phonological and phonetic theories and descriptions are increasingly being operationally tested by the synthesis of their predictions, using speech synthesis systems (Hertz 1999; Dirksen & al. 1997; Bachan & al. 2006). Operational tests are necessary though not sufficient criteria for truth, of course, but within the framework of the full Ecological Cycle sketched in the present contribution, this technology has an important role to play.

Correspondingly, it is increasingly being recognised that the linguistic and phonetic sciences cannot be seriously practised without such tools, and formally well-defined resources to which these tools can be applied, in the new paradigm of Documentary Linguistics (Bird & Simons 2003; Gibbon 2003), within which the principle of the Ecological Cycle is strongly represented: the language communities from which data are gathered are given direct and principled ‘payback’ in various forms, based on the documentation constructed from the data, whether as dictionaries or practical speech synthesis systems or in other forms.

Other areas, such as speech recognition, information extraction (including Google-type searching) and machine translation are not only heavily informed by linguistic theories, but also feed back into theory formation by virtue of providing very extensive, heterogeneous and statistically interesting speech and text data which stimulate the examination of new domains and more details in known domains. Linguistics has traditionally handled phonemic, morphemic, word and sentence units and has stopped there. Recent developments in text theory are being driven by text technology methods which are needed to extract information coherently from hypertexts on the internet. These methods are not only available for us to use thankfully to disseminate our online publications by the provision of standard XML formats, Unicode and search machines for these data types, but are also contributing to the development of linguistic theory in domains above the rank of sentence and in the area of multimodal communication, in areas which have traditionally been called text linguistics.

The universal role claimed here for technology in science appears in the contexts discussed in this paper not as a one-way theory-to-application flow, but as a cyclical interaction in which intellectual potential and technological instrumentality are linked within the overall ecological space within which science - phonology and phonetics no less than other sciences - operates and within which scientists exercise or fail to exercise their involvement with and responsibilities to their intellectual and social environments. Not only physicists, chemists and medics have scientific responsibilities towards the community affected by the technologies based on their discoveries.

But technology is not an end in itself in the present context. Successful (but also unsuccessful) technological applications feed back into the basic sciences which feed them, via the formalisations on which they are based. The comparison of Natural Phonology with Optimality Theory indicates that both of these approaches can benefit from increasing the explicitness of notions such as ‘input’ when it comes to incorporating ‘external evidence’ from language acquisition into the purview of the theory, and that both these approaches can benefit from each other. The same applies to the benefits to be derived from a fully explicit formalisation of both these theories, and the deployment of these formalisations via operational models in practical applications which benefit language users and their communities.

7. References

[1] Bachan, Jolanta & Dafydd Gibbon (2006). Close copy speech synthesis for speech perception testing. *Investigationes Linguisticae* (13), Poznań, 9-24.

[2] Barlow, Jessica A. (2001). Case Study: Optimality Theory and the Assessment and Treatment of Phonological Disorders. *Language, Speech, and Hearing Services in Schools* Vol.32 242-256 October 2001.

[3] Beesley, Kenneth R. & Lauri Karttunen (2003). *Finite State Morphology*. Stanford: CSLI Publications.

[4] Belz, Anja (2000). Computational Learning of Finite-State Models for Natural Language Processing. Doctoral thesis, University of Sussex.

[5] Bromberger, Sylvain & Morris Halle (1989). Why Phonology is Different. *Linguistic Inquiry* 20:51-70.

[6] Carson-Berndsen, Julie (1998). *Time Map Phonology: Finite State Models and Event Logics in Speech Recognition*. Dordrecht, Holland: Kluwer Academic Publishers.

[7] Clements, G.N. (1981). The hierarchical representation of tone features. *Harvard Studies in Phonology* Vol. II. G. Clements, ed., 50-107. Reproduced and distributed by Indiana University Linguistics Club.

[8] Dirksen, Arthur & John S. Coleman (1997). All-Prosodic Synthesis Architecture. J. P. H. van Santen, R. W. Sproat, J. P. Olive and J. Hirschberg, eds. *Progress in Speech Synthesis*. New York: Springer-Verlag, 91-108.

[9] Dziubalska-Kołaczyk Katarzyna (2002). *Beats-and-binding phonology*. Frankfurt/Main: Peter Lang Verlag

[10] Gibbon, Dafydd (1987). Finite state processing of tone systems. *Proc. 3rd European Association of Computational Linguistics Conference* 1987, 291-297.

[11] Gibbon, Dafydd, Roger Moore & Richard Winski (1997). *Handbook of Standards and Resources for Spoken Language Systems*. Berlin: Mouton de Gruyter.

[12] Gibbon, Dafydd (2001). Finite state prosodic analysis of African corpus resources. *EUROSPEECH-2001*, 83-86.

[13] Gibbon, Dafydd (2002). Workable Efficient Language Documentation: a Report and a Vision. *ElsNews* 11.3, 3-5.

[14] Gibbon, Dafydd (2004). Two problems and two methods for prosodic typology. *Proc. Int. Symposium on Tonal Aspects of Languages with Emphasis on Tone Languages*, Beijing.

[15] Gibbon, Dafydd (2005). Forms and formalisation. An essay on formal linguistics. In: Katarzyna Dziubalska-Kołczyk, Jacek Witkoś, Gzegorz Michalski, Bartosz Wiland, eds., *Proceedings of the 1st Student Conference on Formal Linguistics*. Poznan: Adam-Mickiewicz University. Pp. 3-27.

[16] Gibbon, Dafydd, Eno-Abasi Urua, Moses Ekpenyong (2006). Problems and Solutions in African Tone Language Text–To–Speech. In: *Proc. ISCA Tutorial and Research Workshop on Multilingual Speech and Language processing*, Stellenbosch, South Africa.

[17] Govender, Natasha, Etienne Barnard, Marelie Davel (2005). Fundamental Frequency and Tone in isiZulu: Initial Experiments. *Proc. Interspeech2005 - Eurospeech*, Lisbon, 1417-1420.

[18] Hertz, Susan R., Rebecca J. Younes & Nina Zinovieva (1999). Language universal and language specific components in the multi-language ETI Eloquence Text-To-Speech System. *Proc. 14th ICPhS, San Francisco*, 2283-2286.

[19] Hyman, Larry M. (in press). Kuki-Thaadow: An African Tone System in Southeast Asia. In Franck Floricic (ed.), *Mélanges offerts à Denis Creissels*. Les Presses de l'Ecole Normale Supérieure.

[20] Jansche, Martin (1968). A two-level take on Tianjin Tone. In *Proc. 10th European Summer School in Logic, Language and Information, Student Session*. Saarbrücken, 162-174.

[21] Kaplan, Ronald M. & Martin Kay (1994). Regular models of phonological rule systems. *Computational Linguistics* 20(3): 331-378

[22] Karttunen, Lauri (1998). *The Proper Treatment of Optimality in Computational Phonology*. In: *Proceedings of FSMNLP'98. International Workshop on Finite-State Methods in Natural Language Processing, June 29-July 1, 1998*, Bilkent University, Ankara, Turkey, pp. 1-12. Also: http://roa.rutgers.edu/files/258-0498/roa-258-karttunen-2.pdf.

[23] Karttunen, Lauri (2006). The Insufficiency of Paper-and-Pencil Linguistics: the Case of Finnish Prosody. In: Miriam Butt, Mary Dalrymple, and Tracy Holloway King (eds), *Intelligent Linguistic Architectures: Variations on themes by Ronald M. Kaplan*. Stanford, CA: CSLI Publications, pp. 287-300.

[24] Koskenniemi, Kimmo (1983). *Two-level Morphology: A General Computational Model for Word-Form Recognition and Production,* Publications. University of Helsinki, Department of General Linguistics.

[25] Liberman, Mark & Janet B. Pierrehumbert (1984). Intonational invariance under changes in pitch range and length. Aronoff, Mark & Richard T. Oehrle, eds., *Language, sound, structure*. Cambridge: MIT Press, 157-233.

[26] Tucker, Roger & Ksenia Shalonova (2005). Supporting the Creation of TTS for Local Language Voice Information Systems. In: P*roc. Interspeech'2005 - Eurospeech*, Lisbon, 453-456.

[27] Urua, Eno-Abasi, Gibbon, Dafydd & Gut, Ulrike (2003). A Computational Model of Low Tones in Ibibio. *Proceedings of the ICPhS conference 2003*, Barcelona, 623-626.

[28] Vennemann, Theo (1988), *Preference Laws for Syllable Structure*, Berlin: Mouton de Gruyter.

[29] Winograd, Terry (1971). *Procedures as a representation for data in a computer program for understanding natural language*. MIT dissertation. ftp://publications.ai.mit.edu/ai-publications/pdf/AITR-235.pdf (consulted 2009-03-08).