

FINITE STATE PROCESSING OF TONE SYSTEMS

Dafydd Gibbon

(U Bielefeld)

ABSTRACT

It is suggested in this paper that two-level morphology theory (Kay, Koskeniemi) can be extended to include morphological tone. This extension treats phonological features as I/O tapes for Finite State Transducers in a parallel sequential incrementation (PSI) architecture; phonological processes (e.g. assimilation) are seen as variants of an elementary unification operation over feature tapes (linear unification phonology, LUP). The phenomena analysed are tone terracing with tone-spreading (horizontal assimilation), downstep, upstep, downdrift, upsweep in two West African languages, Tem (Togo) and Baule (Côte d'Ivoire). It is shown that an FST account leads to more insightful definitions of the basic phenomena than other approaches (e.g. phonological rules or metrical systems).

1. Descriptive context

The topic of this paper is tone sandhi in two West African tone languages and suitable formal models for it. The languages investigated are Tem (Gur/Voltaic family, Togo) and Baule (Akan family, Côte d'Ivoire). Tone languages of other types, in particular the Sino-Tibetan languages, will not be discussed.

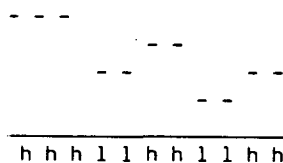
The specific concern of this paper is with the way in which certain quite well-known morpho-phonological (lexical) tone patterns are realized in sequence in terms of phonetic pitch patterns. There are three interacting factors involved: i. tone-text association rules; ii. tone-sandhi rules; iii. phonetic interpretation rules.

Tone-text association rules are concerned with the association of tones with syllables (primary associations and a form of tone spreading) as well

as floating tones and compound tones. Floating tones are not associated with syllables, but are postulated to explain apparent irregularities in phonetic patterning in terms of regular tone sandhi properties.

The tone sandhi rules define how tones affect their neighbours. The example to be treated here is a kind of tonal assimilation known as tonal spreading in which low tones are phonetically raised following a high tone or, more frequently, high tones are lowered after a low tone, either to the level of the low tone (total downstep) or to a mid level (partial downstep). The newly defined tone is then the reference point for following tones.

The latter kind of assimilation produces a characteristic perceptual, and experimentally measurable, effect known as tone terracing. Tone sequences are realized at a fairly high level at the beginning of a sequence, and at certain well-defined points the whole pitch register appears to be downstepped to a new level. The process may be iterated several times. It is often represented in the literature in the following way (partial downstep); it can be seen that a later high tone may be as high as or lower than an earlier low tone:



In particular, it will be seen that the two terraced tone languages, Tem and Baule, involve similar processes in detail and have similar basic FST architectures, but differ systematically at certain well-defined points involving sandhi generality, and scope of sandhi context.

Detailed phonetic interpretation involves pitch patterns between neighbouring tones of the same type within terraces. These are processes of downdrift (neighbouring tones fall) or upsweep (neighbouring tones, usually high tones, rise). They will not be dealt with here.

2. Theoretical context

The view is developing, based on work by Kay and Kaplan, Koskenniemi, Church, and others, that in phonology it is sufficient to use finite state transducers which are not allowed to apply to their own output. Kay and Kaplan have shown that it is possible to reduce conventional, so-called "context-sensitive" phonological rules to finite-state relations, and to apply the FSTs thus produced in sequential order (Kay 1986).

Koskenniemi developed a somewhat different concept for Finnish morphology, in which the FSTs operate as parallel filters over the input: they must all agree in their output. A careful analysis also shows that Church's allophonic parser, in his actual implementation using matrix operations to simulate bottom up chart parsing, can also be seen as a system of parallel finite state filters. The PSI (Parallel Sequential Incrementation) system of prosodic analysis being developed by myself in the DFG Forschergruppe "Kohärenz" in Bielefeld incorporates a similar concept of FSTs used as a parallel filter bank (Gibbon & al. 1986).

The context within descriptive phonology is that of theories which postulate interrelated but structurally autonomous parallel levels of organization in phonology. The four major classical directions in this area are traditional intonation analysis (surveyed and developed in Gibbon 1976), Firthian "prosodic phonology", Pike's simultaneous hierarchies, and the non-linear (autosegmental and metrical) phonologies of the last thirteen years.

Parallel FST systems are used in order to explicate both traditional phonological rules, so long as they do not apply to their own output, and also, with appropriate synchronization measures, the parallel tiers which figure in autosegmental phonology, the mappings between these tiers, and the

mappings between abstract and concrete phonological and phonetic levels of representation. FST systems are conceptually bidirectional; they may easily be processed in either direction with the same finite state mechanism; the problem of the recoverability of underlying structure (short of ambiguity through genuine neutralization) loses its importance.

The idea of formulating prosodic patterns in English intonation in FS terms was originated and developed by Pierrehumbert (1980), though FS intonation models had been developed much earlier by 't Hart and others for Dutch intonation. These existing FS intonation descriptions are straightforward finite state automata (FSAs; for Dutch, probabilistic FSAs). The problem of mapping such patterns at one level on to patterns at another, the traditional problem in descriptive linguistics as well as in computational parsing and translation, has not been formulated in finite state terms for this domain. This mapping question is a different one from the question of recognition, and the finite state devices required for an answer to the question are different. Additionally, the tone language application constitutes a different domain.

The input and output languages for FSTs are both regular sets (Type 3 languages). FSTs have various interesting properties which are in part similar to those of FSAs. The reversibility property shown in the simulations is one of the most interesting. Any FST which is deterministic in one direction is not necessarily deterministic in the other, as the neutralization facts in Tem and Baule show. Furthermore, it is not true for FSTs, as it is for FSAs, that for any non-deterministic FST there is a deterministic one which is weakly equivalent relative to the input language. This only holds if the paired input and output symbols are interpreted as compound (relational) input symbols, and the input and output tapes are seen as a single tape of pairs. This is an abstraction which formally reduces FSTs to FSAs. Kay has suggested this perspective on FSTs as an explication for relations between linguistic levels, where FSTs define relations between linguistic levels of representation in an essentially declarative fashion, though with a procedural interpretation. For a slightly different FST definition cf. Aho & Ullman (1972). In current computational theories of language (FUG, GPSG,

LFG), the standard treatment for concord restriction, to which phonological assimilation and neutralization may be compared, is in terms of a class of operations related to unification. The situation in autosegmental phonology is simpler than in syntax, in that each feature or tier can be modelled by a finite state device. The elementary unification operator required is, correspondingly, restricted to non-recursive, adjacent feature specification on a given tier, as in the present analysis. In a non-parallel architecture, the operation would be embedded in a more complex, perhaps context-free-style, context.

3. The Tem and Baule data

The essential tonal properties of Tem are: downstep, downdrift, phonetically constant (non-terraced) low tone, high tone spreading over a following low; only terracing and sandhi are dealt with here. The Tem data and the inter-level relations are taken from Tchagbale (1984). The following shows a simulation using a bidirectional FST interpreter, with runs in each direction.

Forward:

```

IN1      (L L L L)
OUT1     1 (LC LC LC LC)

IN1      (H H H)
OUT1     1 (HC H H)

IN1      (L H H)
OUT1     1 (LC !H H)

IN1      (L H L L)
OUT1     1 (LC !H H LC)

IN1      (L H *L)
OUT1     1 (LC !H LC)

IN1      (L L L H)
OUT1     1 (LC LC LC !H)

IN1      (L H L H)
OUT1     1 (LC !H H !H)

IN1      (L H L L H)
OUT1     1 (LC !H H LC !H)

IN1      (L H L H H)
OUT1     1 (LC !H H !H H)

IN1      (L H L L H L L)
OUT1     1 (LC !H H LC !H H LC)

IN1      (L H L H H H)
OUT1     1 (LC !H H !H H H)

IN1      (H L H L L)
OUT1     1 (HC H !H H LC)

```

Reverse:

```

IN1      (LC LC LC LC)
OUT1     1 (L L L L)

IN1      (HC H H)
OUT1     1 (H H H)
OUT2     2 (H H L)

IN1      (LC !H H)
OUT1     1 (L H H)
OUT2     2 (L H L)

IN1      (LC !H H LC)
OUT1     1 (L H H *L)
OUT2     2 (L H L L)

IN1      (LC !H LC)
OUT1     1 (L H *L)

IN1      (LC LC LC !H)
OUT1     1 (L L L H)

IN1      (LC !H H !H)
OUT1     1 (L H L H)

IN1      (LC !H H LC !H)
OUT1     1 (L H L L H)

IN1      (LC !H H !H H)
OUT1     1 (L H L H H)
OUT2     2 (L H L H L)

IN1      (LC !H H LC !H H LC)
OUT1     1 (L H L L H H *L)
OUT2     2 (L H L L H L L)

IN1      (LC !H H !H H H)
OUT1     1 (L H L H H H)
OUT2     2 (L H L H H L)

IN1      (HC H !H H LC)
OUT1     1 (H L H H *L)
OUT2     2 (H L H L L)

```

The essential tonal properties of Baule are: partial or total downstep (style-determined), up-step, upsweep, downdrift, tone spreading of both low and high over the first tone of an appositely specified sequence, compound tone. Again, only terracing and sandhi are dealt with. The Baule sandhi data are from Ahoua (1987a), simulated by the same interpreter, with an FST designed for Baule.

Forward:

IN_s (H L L L L)
 OUT_s 1 (HC H L L L)

IN_s (H L L)
 OUT_s 1 (HC H L)

IN_s (L H L L)
 OUT_s 1 (LC !H H L)

IN_s (L L L H L L L)
 OUT_s 1 (LC L L !H H L L)

IN_s (L H H H H)
 OUT_s 1 (LC L !H H H)

IN_s (H L H H H H)
 OUT_s 1 (HC L L !H H H)

IN_s (L L L L H L)
 OUT_s 1 (LC L L L !H L)

IN_s (L H H)
 OUT_s 1 (LC L !H)

IN_s (L H L H)
 OUT_s 1 (LC !H L !H)

IN_s (L H L H L)
 OUT_s 1 (LC !H L !H L)

IN_s (L H L L H)
 OUT_s 1 (LC !H H L !H)

IN_s (L H)
 OUT_s 1 (LC !H)

IN_s (H H)
 OUT_s 1 (HC H)

IN_s (L L)
 OUT_s 1 (LC L)

IN_s (H L)
 OUT_s 1 (HC L)

Reverse:

IN_s (HC H L L L)
 OUT_s 1 (H L L L L)

IN_s (HC H L)
 OUT_s 1 (H H L)
 2 (H L L)

IN_s (LC !H H L)
 OUT_s 1 (L H L L)

IN_s (LC L L !H H L L)
 OUT_s 1 (L L L H L L L)
 2 (L L H H L L L)

IN_s (LC L !H H H)
 OUT_s 1 (L H H H H)

IN_s (HC L L !H H H)
 OUT_s 1 (H L H H H H)

IN_s (LC L L L !H L)
 OUT_s 1 (L L L L H L)
 2 (L L L H H L)

IN_s (LC L !H)
 OUT_s 1 (L L H)
 2 (L H H)

IN_s (LC !H L !H)
 OUT_s 1 (L H L H)

IN_s (LC !H L !H L)
 OUT_s 1 (L H L H L)

IN_s (LC !H H L !H)
 OUT_s 1 (L H L L H)

IN_s (LC !H)
 OUT_s 1 (L H)

IN_s (HC H)
 OUT_s 1 (H H)

IN_s (LC L)
 OUT_s 1 (L L)

IN_s (HC L)
 OUT_s 1 (H L)

The underlying morphophonological tones are annotated as follows:

L = low

H = high

*L = low with an additional morphological feature (Tem only).

The surface phonetic tones are:

LC = low constant (in Baule, only initial)

HC = high constant (only initial)

H = high relative to currently defined level

L = low relative to currently defined level
(Baule)

!H = mid (=downstepped high) tone.

The simulations show the properties of the tone sandhi systems of Tem and Baule very clearly, in particular the contextual dependencies (sandhi). The reverse (recognition) simulations show the effects of tone neutralization: in the reverse direction, non-deterministic analyses are required, which means in the present context that more than one underlying form may be found.

The tone systems of Tem and Baule can be seen to differ in several important respects, which are summarized in the transition network representations given in Figures 1 and 2, respectively.

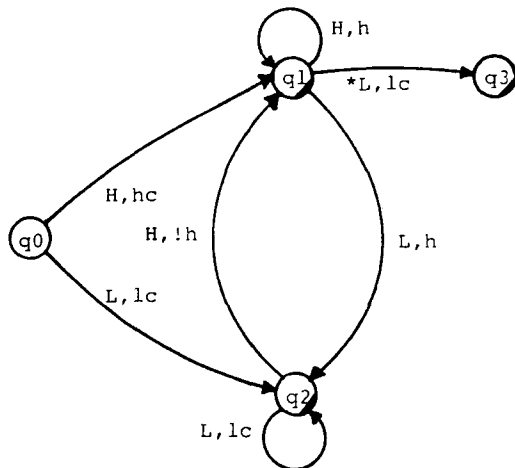


Figure 1: The Tem FST

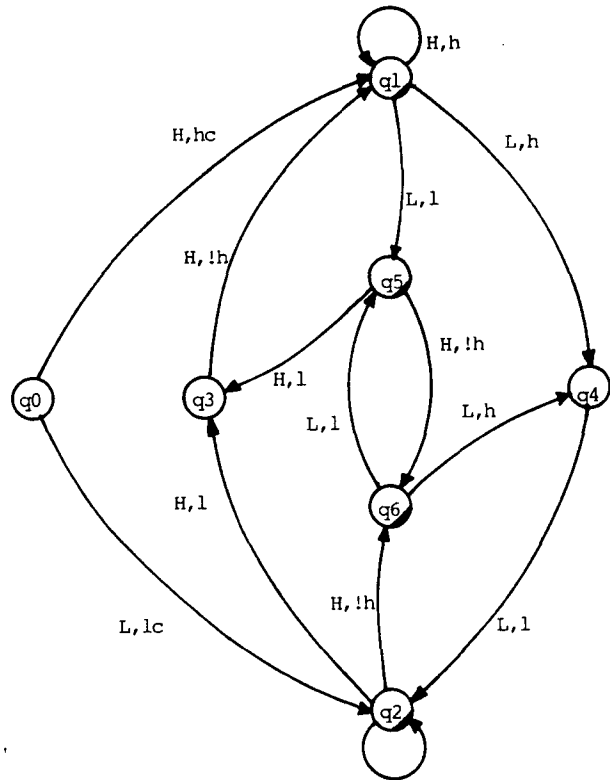


Figure 2: The Baule FST

Another interesting point pertains to local vs. global pitch relations. The relations described here are clearly local, if they are formalizable in finite state terms. But this is not to say that there is not a global factor involved in addition to these local factors. On the contrary, Ahoua(1987b) has demonstrated the presence of global intonational factors in Baule which are superimposed on the local tonal relations and partly suppress them in fast speech styles.

4. Conclusion

It is immediately obvious that the transition diagramme representations show similar iterative cyclical processes for Tem and for Baule; the Baule system has an "inner" and an "outer" cycle, which may be accessed and left at well-defined points. At corresponding points in the diagrammes,

both systems show "epicycles", i.e. transitions which start and end at the same node, and the tone assimilation transitions also occur at similar points in the systems relative to the epicycles.

The suggested interpretation for these interrelated iterative process types, three in Tem and five in Baule, is that they are immediately plausible explications for the concept of linguistic rhythm and interlocking rhythmic patternings. This is the same explicandum, fundamentally, as in metrical phonology, but it is associated here with the claim that an explicit concept of iteration is a more adequate explanation for rhythm than a tree-based, implicitly context-free notation, which is not only over-powerful but also ill-suited to the problem, or traditional phonological rules, whose formal properties are unclear.

The formal properties of Tem and Baule as terraced tone languages can be defined in terms of the topology of the FST transition diagrams:

- i. The fundamental notion of "terrace" or "tonal unit" is defined as one cycle (iteration, oscillation) between major nodes of the system.
- ii. A major node is a node which has unlike input symbols on non-epicyclic input and output transitions and can also be a final node.
- iii. Terrace-internal monotone sequences are defined as epicycles; in Baule, epicyclic sequences start not on the second but on the third item of the sequence, and a non-epicyclic sub-system is required.
- iv. Stepping and spreading occurs on any non-epicyclic transition leaving a major node; in languages with downstep only (Tem), this only applies to high tones, in those with downstep and upstep, upstep occurs with low tones in these positions.

These definitions show that the FST formalism is not just another "notational variant", but precise, highly suggestive, and useful in that it is a formally clear and simple system with well-understood computational properties which make it easy to implement tools for testing the consistency and completeness of a given description.

In current non-linear approaches in descriptive phonology it is not clear that the basic explicanda—types of iteration or rhythm, the character of terracing as a particular kind of iteration or oscillation, and the relative complexity of different tone systems — are captured by the notation, in contrast to the clarity and immediate interpretability of the FST model. In one current model (Clements, communicated by Ahoua), complex constructive definitions are given; they may be characterized in terms of conventional parsing techniques as follows:

- i. analyze the input string into "islands" which define the borders between tone terraces;
- ii. proceed "bottom up" to make constituents (feet, in general to the left) of these islands;
- iii. proceed either bottom up or top down to create a right-branching tree over these constituents.
- iv. (implicit) perform tonal assimilation on the left-most tone on each left branch.

This is an unnecessarily complex system, whose formal properties (context-free? bottom up? right-left?) are not clear.

A complete evaluation of different approaches will clearly require prior elaboration of the tone-text association rules and the phonetic interpretation rules. The former will follow the principles laid down in Goldsmith's well-formedness condition on tone alignment, which also point to the applicability of FST systems.

In summary, the prospects for a comprehensive FST based account of morphophonological tone phenomena appear to be good. The prospects are all the more interesting in view of the developments in FS morphology and phonology over the past four years, suggesting that an overall model for all aspects of sublexical processing may be feasible using an overall parallel sequential incrementation (PSI) architecture with FST components for inter-level mapping. It may be predicted with some hope of success that components which are

more powerful than Finite State will turn out to be unnecessary, at least for the sublexical domain, even outside the conventional area of Western European languages.

Tchagbale, Z. 1984. T.D. de Linguistique: exercices et corrigés. Institut de Linguistique Appliquée, Université Nationale de Côte-d'Ivoire, Abidjan, No. 103.

References

- Aho, A.V., J.D. Ullman 1972 **The Theory of Parsing, Translation, and Compiling**. Vol.1: Parsing. Prentice-Hall, Englewood Cliffs, N.J.
- Ahoua, F. 1987a. "Government in West African tonal systems with special reference to Baule and Dafing." To appear.
- Ahoua, F. 1987b. "Tone and Intonation in Baule." Paper held at the DGfS Annual Conference, Augsburg.
- Church, K.W. 1980. On Memory Limitations in Natural Language Processing. Distributed by IULC, 1982.
- Church, K.W. 1983. Phrase-Structure Parsing: A Method for Taking Advantage of Allophonic Constraints. Dissertation, MIT.
- Clements, G.N. 1981. "The hierarchical representation of tone features." Harvard Studies in Phonology 2. Distributed by IULC.
- Gibbon, D. 1976. Perspectives of Intonation Analysis. Berne, Lang.
- Gibbon, D., G. Braun, F. Jin, V. Pignataro 1986. Prosodic Cohesion. Interim Project Report, DFG-Forschergruppe "Kohärenz", U Bielefeld.
- 't Hart, J. & R. Collier. 1975. "Integrating different levels of intonation analysis." **Journal of Phonetics** 3: 235-255.
- Kay, M. 1986. Lectures on Unification Grammar. DGfS Summer School, Munich.
- Kaplan, R. & M. Kay. 1981. "Phonological rules and finite-state transducers." Paper at the Annual Meeting of the ACL, 28.12.1981. NYC. (Cited by Koskenniemi.)
- Koskenniemi, K. 1983. Two-level Morphology: A General Computational Model for Word-Form Recognition and Production. Dissertation, U Helsinki.
- Marcus, M. 1980. **A Theory of Syntactic Recognition for Natural Language**. MIT Press, Cambridge, Mass.
- Pierrehumbert, J.B. 1980. The Phonology and Phonetics of English Intonation. Diss. MIT.