A Formal Functional Model of Tone

Building on the insights of Tsay (1994) and the functional Optimality Theory formalism of Boersma (1998), this paper argues for a model of tone that attempts to explain how cross-linguistic generalizations arise from extragrammatical causes. We hope that careful study of this phonetically simple subarea of phonology will make it easier to judge the feasibility of formal functional approaches to phonology in general.

The central claim in functional OT is that phonological systems involve separate production and perception grammars that are coordinated during acquisition. The grammars consist of ranked constraints of high phonetic detail; categoricality and contrastiveness are not inherent but rather emerge through constraint ranking (see also Kirchner 1997). For example, figure (1) illustrates how a system with two level tones emerges from the ranking of *CATEG(F) ("perceptual feature-value F is not used as a phonological category") and *WARP(D) ("a perceptual feature-value cannot be distorted by amount D from the acoustic feature-value") (cf. Boersma 1998:165; numbers = F0 as percentage of listener's pitch range, "[]" = acoustic forms, "//" = perceptual forms). Additional functional restrictions are formalized via universal ranking principles, e.g. minimization of distortion (i.e. *WARP(D1) >> *WARP(D2) iff D1 > D2) and category finiteness (i.e. *CATEG(F) is undominated except for a finite set of F values) (Boersma 1998). Our task is to motivate further independently motivated ranking principles that capture cross-linguistic generalizations about tone systems.

The first generalization we address is that there are fewer languages with n contrasting tone levels than n-1 tone levels (n \geq 2), with no attested system having more than five contrastive levels (Maddieson 1978). The first part of this generalization follows from category finiteness if we assume that positive evidence is required to demote a given *CATEG constraint during acquisition; each new demotion adds to the overall cost in acquiring the system. The five-level limit is not absolute, but is strongly motivated by limits on working memory capacity (Tsay 1994). Formally, this is encoded by an acquisition principle making it costly to demote *CATEG below *WARP with distortion near a memory-motivated threshold (e.g. disfavoring the ranking of *CATEG below *WARP(20) would disfavor tone systems with more than five levels). By contrast, tone theories that posit innate tone features (Yip 1995) do not predict the existence of five-level systems and/or do not explain why even four-level systems are significantly rarer than three-level systems.

The model also predicts that tonal phenomena with different teleologies (articulatory vs. perceptual) will behave differently. Tone spreading, for example, is generally agreed to arise through the phonologization of coarticulation. In this model this cause is formalized in the production grammar (via a principle giving a default high ranking to constraints of the form *CHANGE, which disallow changing articulatory configurations). The model thus correctly predicts that phonological tone spreading phenomena should respect articulatory constraints on tone production. For example, the spreading of tone may be blocked by obstruents with incompatible laryngeal properties (e.g. Bade, Nupe [Odden 1995:451-2]). It also appears that register alone may sometimes spread (e.g. Ewe [Odden 1995:453], Pingyao [Yip 1995:483-4]), where upper vs. lower register are articulated by distinct muscle groups (Ohala 1978). However, tone contour shape cannot spread (Yip 1995:480-1), since the model provides no mechanism for the spread of sequences of articulatory gestures; if contours are allowed by the production grammar, the relevant *CHANGE constraints must have been demoted, and so cannot force spreading.

By contrast, local dissimilation has a perceptual rather than articulatory teleology, arising when listeners perceive an intended articulatory gesture as mere coarticulation and remove it from the representation (Ohala 1986). This can be formalized in the present model by ranking constraints in the perception grammar to allow for greater tolerance of
distortion of a tone in the context of another tone. Given this, the model predicts that the production grammar should be able to target classes of tones produced with distinct articulatory gestures as long as they are acoustically similar. For example, in the four-level tone system of Nikki Bariba (Welmers 1952), the two mid tones behave as a class in a dissimilation process because they are perceptually similar, even though articulatorily this class crosses the register boundary (see figure (2)). Domain-limit rules (Nespor and Vogel 1986) also have a perceptual teleology since they enhance the salience of boundaries; they are thus also correctly predicted to ignore register (e.g. tone raising in Igede [Bergman 1971], tone sandhi in Taiwanese [Tsay 1994]). Standard theories of tone features do not predict this relationship between rule type and the relevance/irrelevance of laryngeal features.

Figures.

(1) [50] *CATEG(90) *WARP(20) *CATEG(30) *CATEG(60)

<table>
<thead>
<tr>
<th>[30]</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>/30/</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>/60/</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/90/</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

(2) Both mid tones (2, 3) are raised one degree (to 3, 4 respectively) before lowest tone (1)

go\textsuperscript{1}na\textsuperscript{2} "guinea fowl"  na\textsuperscript{3} ta\textsuperscript{3} su\textsuperscript{3} duu\textsuperscript{2} ra\textsuperscript{2} "I planted yams"
goa\textsuperscript{1}na\textsuperscript{3} ye\textsuperscript{1} ni\textsuperscript{4} "that guinea fowl over there"  na\textsuperscript{4} ra\textsuperscript{1} ta\textsuperscript{3} su\textsuperscript{4} duu\textsuperscript{1} re\textsuperscript{1} "I plant yams"

References.