

A Formal Functional Model of Tone^{*}

James Myers and Jane Tsay
National Chung Cheng University

Formalist and functionalist approaches to language have long been seen as inherently antagonistic, but in recent years this has begun to change, as formalists recognize the influence of language use on linguistic structure, and functionalists provide more explicit models for how this influence operates. The model of tone proposed here is formal in the sense of aiming to provide a completely explicit account of phonological knowledge, yet it is also functionalist, since its elements take the form they do because of the way phonological knowledge is actually used when speaking, hearing, and acquiring a language. In particular, this paper builds on the tonal insights of Tsay (1994) and the functional Optimality Theory formalism of Boersma (1998) to propose a new model that explains important aspects of tonal behavior: limits on the number of level tones, interactions of tones and other features, and tone spreading. Data primarily come from tone languages of the Sino-Tibetan family.

Key words: phonology, tone, functionalism, Optimality Theory

1. Introduction

Formalist and functionalist approaches to language have long been seen as inherently antagonistic. However, in recent years this has begun to change, as formalists recognize the influence of language use on linguistic structure, and functionalists provide more explicit models for how this influence operates. We intend this paper to be a part of this growing trend. The model of tone proposed here is formal in the sense of aiming to provide a completely explicit account of phonological knowledge, and yet the model is also functionalist, since its elements take the form they do because of the way phonological knowledge is actually used when speaking, hearing, and acquiring a language.

For our larger purposes, tonal phenomena have the advantages of being both

^{*} This is a revised and shortened version of a paper first presented at GLOW in Asia, held at National Tsing Hua University, Taiwan, in January, 2002. We would like to thank GLOW attendees, two anonymous reviewers, and Paul Boersma for insightful comments that we have done our best to incorporate. Of course we are solely responsible for any errors in data description or analysis that may remain.

phonologically quite intricate and yet phonetically relatively straightforward (i.e., involving primarily a single perceptual dimension, although laryngeal physiology is admittedly more complex). An additional advantage is that a detailed functionalist model of tone is already available, namely Tsay (1994). However, although its goals were functionalist in the sense we mean here, the model presented in Tsay (1994) was hampered by the formal devices available at the time, with the result that the most interesting tonal phenomena were made the responsibility of the least explicit parts of her theory.

However, there has since developed a theoretical framework that makes it far easier to give functional models formal rigor: Optimality Theory (Prince and Smolensky 1993, henceforth OT). While most of the OT literature is not explicitly labeled as such, some recent works wear the functionalist label with pride, in particular the theory of Functional Phonology presented in Boersma (1998) (see also Flemming 1995, Kirchner 1997, 1998, Hayes 1999, Zhang 2002, and the works introduced by Gussenhoven and Kager 2001, among many others).¹ Functionalist OT models are formal in being completely explicit in how constraints are selected, ranked, and evaluated, and yet they are functionalist in that the constraints (and to a certain extent their rankings) reflect extra-grammatical factors such as speech physiology and auditory perception.

This paper builds on the functionalist insights into tone of Tsay (1994) and the formalism of Boersma (1998) to argue for a new model that attempts to describe how cross-linguistic tonal generalizations arise from extra-grammatical causes, supporting the model with data from tone languages primarily of the Sino-Tibetan family. This model thus takes quite a different approach from more traditionally formalist analyses of tone (most recently exemplified by Bao 1999, Chen 2000, and Yip 2002), but we are not the first to attempt a functionalist analysis of tone within OT formalism: Zhang (2002), an in-depth phonetically motivated analysis of contour tone distribution, is perhaps the most illustrious of our predecessors.

We begin in section 2 by reviewing the philosophical and formal underpinnings of the model. Next we apply the model in analyses of some fundamental aspects of tonal phenomena: section 3 is concerned with the nature of tone features, while section 4 examines the autosegmental behavior of tone. Finally, section 5 provides a summary of our major conclusions and the work that remains to be done.

¹ There is also a growing literature on functionalist syntax within an OT framework (see, e.g., Legendre, Grimshaw, and Vikner 2001, Sells 2001).

2. Functional Phonology

We begin in this section with further clarification of the notion of a formal functional model of phonology. In 2.1 we explain why a purely formal approach to phonology cannot work, and why we think OT permits a reasonable formalism for functionalist phonology. Then in 2.2 we give an introduction to the specific model of functionalist OT that we use in this paper, the Functional Phonology model of Boersma (1998).

2.1 Formalism and functionalism in phonology

By its very nature, phonology acts as an interface between mental representations and physical signals. The question of how to balance formalism and functionalism in phonological theory is thus an old one; already in the primary text of generative phonology, the arch-formalists Chomsky and Halle (1968:400) admitted that their approach was “overly formal” since it ignored the “intrinsic content” of phonological symbols. While all generative phonologists since then have recognized the relevance of the physical aspects of speech, they have varied widely in where they draw the line between “mere phonetics” and “true phonology.” At one extreme are approaches like that of Hale and Reiss (2000), who reduce “true” phonological knowledge to an extremely restricted formal minimum, a formal core designed to capture all “*computationally possible* human grammars” (Hale and Reiss 2000:167, *italics in original*), not just the ones that are possible to learn or capable of developing through diachronic processes like the phonologization of phonetics.

The major trouble with this approach is that a model that is so abstract that it makes no reference to phonetics cannot answer most of the questions that have occupied phonological theory for the past several decades (e.g., the phonological behavior of features). Theorists like Hale and Reiss (2000) pass such questions off to extra-grammatical factors, like phonetics and acquisition, but they intentionally leave these unformalized, since computation is claimed to involve only abstract symbols and rules. With the most explanatory part of their model left unformalized, it is not clear whether it can be called truly “generative” in the sense of Chomsky (1965:4), that is, completely explicit. Even if it could be shown to be wise to use Occam’s razor as they do, cutting virtually all of traditional phonology out of phonological theory, the interface role of phonology means that the choices about what belongs within the formalism and what belongs outside it become quite arbitrary.

Considerations such as these have led us to conclude that the best approach to the problem of formalism and functionalism in phonology is to stretch the formal boundaries in the opposite direction, as widely as possible. If the psychophysics of

articulation or perception are relevant for explaining a phonological pattern, then phonologists should make an effort to incorporate psychophysics into the formal model. Not only will this make phonological theory more explicit, but it will also allow the theory to capture the fundamental nature of phonology as an interface system.

An important implication of this functional approach is that it puts phonology on the same branch of cognitive science as vision and motor control, whose primary job is to develop and use mental representations of the physical world. That is, phonological knowledge is precisely the knowledge of (one kind of) physics, perhaps supplemented with innate biases specific to the phonological processing systems of the brain. This view of phonology as a branch of psychophysics is supported by discoveries made over the last few decades that much of what had been dismissed as “pure physics” in speech is in fact under systematic (and learned) psychological control. Prominent examples of this include the variability of subphonemic vowel lengthening effects across languages (Chen 1970) and systematic variations in the articulation of /l/ in English (Sproat and Fujimura 1993); see Keating (1985, 1990) and Port (1996) for reviews of many other cases. At the very least, such discoveries demonstrate that speakers’ knowledge of the sound system of their language (what is usually called “phonology”) goes far beyond the abstract categorical confines assumed by much of phonological theory.

Only recently, however, has generative theory developed a formalism truly capable of treating phonology in the functional manner necessitated by its interface nature. This is the formalism of Optimality Theory (OT), whose advantages for functionalist modeling are also noted by Hale and Reiss (2000): OT constraints formalize the notion of markedness, long a staple of functionalist approaches, and the OT device of constraint ranking easily implements the functionalist notion of structure emerging through compromises between equally valid grammar-external forces (e.g., ease of articulation vs. perceptibility). Exploiting concepts like these, functionalist OT models (e.g., Flemming 1995, Kirchner 1997, 1998, Boersma 1998, Hayes 1999, Gussenhoven and Kager 2001, Zhang 2002) build grammars solely out of constraints whose existence and ranking are motivated by psychophysics, forming what Hayes (1999) calls “phonetically driven phonology”. In the most radical of these models, familiar categorical units like distinctive features are given up entirely, since as Kirchner (1997) and Boersma (1998) independently discovered (and as will be reviewed in the next section), key aspects of categoricity can be made to emerge from competing OT constraints operating on representations that are phonetically quite detailed. Hence functional OT treats phonology as knowledge of speech physics, yet without sacrificing formal rigor or the ability to describe categorical phonology as well.

By formalizing functionalism, rather than attempting a strict segregation between phonological patterns and their causes, the functional OT approach seems to us to

provide one of the most promising generative approaches to phonological theory currently available.

2.2 The theory of Functional Phonology

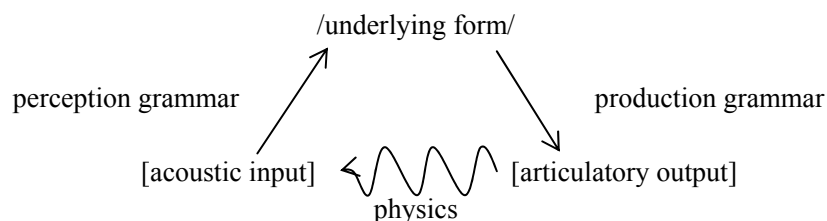
Probably the most explicit and complete functional OT model is the theory of Functional Phonology, presented in Boersma (1998). This model is still undergoing development (see, e.g., Boersma 1999, Boersma and Hayes 2001, Escudero and Boersma 2001) and has a large number of components, not all of which will be employed in this paper. For the purposes of formal explicitness, however, we shall adopt its claims, analyses, and terminology wherever they relate to our own empirical issue, phonological tone. We first provide a detailed introduction to these elements of the theory in 2.2.1, and then in 2.2.2 we address more general issues raised by the theory.

2.2.1 Functional Phonology formalism

The central claim of Boersma (1998) is that phonological systems involve separate grammars for production, perception, and recognition. These grammars divide the familiar speech cycle into three mappings, respectively: between underlying forms and articulatory forms (production grammar), between psychophysical forms (e.g., for spoken language, psychoacoustic forms) and perceptual forms (perception grammar), and between perceptual forms and the underlying forms (recognition grammar). In this paper we shall use a simplified version of the model with only two grammars, a production grammar and a perception grammar (conflating Boersma's perception and recognition grammars). Our perception grammar, therefore, maps between psychophysical inputs and underlying forms, which are thus perceptual targets; these same perceptual targets are also used as the inputs to the production grammar.² The production and perception grammars are OT grammars, defined solely by the ranking of structure constraints that evaluate outputs (articulatory or perceptual) and faithfulness constraints that compare the outputs with their associated inputs (underlying forms in the case of the production grammar, psychophysical forms in the case of the perception grammar). The resulting simplified model appears as follows (for more detailed diagrams see Boersma 1998:143, 1999:1).

² Our simplification, while necessary to keep the scope of this paper within reasonable limits, is not completely unproblematic. In particular, it implies that word recognition is entirely psychophysically driven, ignoring the evidence for articulatory representations (Liberman 1982) and top-down processes (Elman and McClelland 1988). These complexities have no effect on the points made in this paper.

(1) The relation between the production and perception grammars³



The production and perception grammars each have their own structure and faithfulness constraints, all of which are couched in detailed psychophysical terms. Prototypical examples of these constraints are given below, with the first of each pair representing a structure constraint (operating solely on outputs) and the second a faithfulness constraint (comparing inputs and outputs).

(2) Production grammar constraints (after Boersma 1998:152, 176; Boersma 1999:7)⁴

- a. *GESTURE (articulator: gesture / distance, duration, precision, velocity)
Do not let a certain articulator perform a certain gesture, along a certain distance, for a certain duration, and with a certain precision and velocity.
- b. *REPLACE (feature: value₁, value₂ / condition / left-env _ right-env)
Do not replace, for a certain feature, a specified value with a different value, under a certain condition and in the environment between left-env and right-env.

(3) Perception grammar constraints (after Boersma 1998:163-5)

- a. *CATEG (feature: value)
The given perceptual feature cannot be recognized as the given value.

³ For sign languages, of course, “acoustic input” must be replaced by “visual input”.

⁴ We have changed the description of production grammar constraint *REPLACE somewhat, in order to accommodate it to the simplified theory presented here. In the full theory, *REPLACE evaluates the faithfulness of the speaker’s perception of her own production to the underlying form, so the features that it mentions are perceptual features. Here we shall usually be able to get away with making the simplifying assumption that the features are instead the acoustic features automatically generated by articulation.

b. *WARP (feature: distortion)

The perceived value of a feature cannot be different from the psychophysical value by the given amount of distortion.

Like standard OT, the structure constraints assign cost to the elements of phonological content (articulatory gestures in the case of *GESTURE, and perceptual features, e.g., fundamental frequency, in the case of *CATEG), while the faithfulness constraints require identity between input and output (in the production grammar, *REPLACE disallows gestures that would produce psychoacoustic forms different from the perceptual targets listed in the lexicon, while in the perception grammar, *WARP disallows mismatches between psychoacoustic forms and these lexical perceptual targets). Crucially, however, the representations being evaluated are phonetically detailed rather than abstract and discrete.

We now illustrate a simple application of these constraints for the fundamental issue of categoricity. The representations and constraints of Functional Phonology are phonetically detailed, so phonetic discreteness is held to be an emergent property, not inherent. The empirical advantages of this approach are reviewed elsewhere (e.g., Kirchner 1997, Myers 1997, Steriade 2000) and further arguments will be discussed below, when we examine tone features in more depth. Here we merely exemplify the proposal.

Suppose we have a tone system with two contrastive tones, whose perceptual targets may be represented as /1/ and /5/ respectively (where 1 = lowest pitch and 5 = highest, following Chao 1930).⁵ The problem is to formalize the observations that, first, speakers of such languages treat the targets /1/ and /5/ as distinct categories, and second, they perceive all phonetic pitches produced by fellow speakers to fall into one of these two categories, even if they are actually somewhat off, e.g., [2] or [4]. Together these observations represent categorical perception. In this theory, the first observation is handled by the principle of Category Finiteness, which assumes that there are an infinite number of *CATEG constraints, one for each possible value of a category, and that all of them begin as undominated (i.e., the initial state of the system is to have no categories at all). Since positive evidence during learning is required to demote a *CATEG constraint (thus allowing its value to represent a category), the mature system can only have a finite number of categories (see Boersma 1998:164).⁶

⁵ This is another oversimplification, since the actual pitch values for tones depend on many different factors that we shall ignore, including pitch range of the speaker, intonation, and the number of level tones in the system (e.g., the more tones there are, the more spread out the pitch range; Maddieson 1978b).

⁶ As a reviewer points out, in Functional Phonology the notion of categories itself does not

The second observation is the responsibility of a second principle, Minimization of Distortion, which states that “a less distorted recognition is preferred over a more distorted recognition” (Boersma 1998:164). Formally, this is handled by requiring that the ranking of *WARP constraints parallels the degree of distortion (measured in psychoacoustic units) that they forbid.

(4) Minimization of Distortion

$$*\text{WARP}(\text{distortion}_1) \gg *\text{WARP}(\text{distortion}_2) \Leftrightarrow \text{distortion}_1 > \text{distortion}_2$$

In our two-tone system, phonetic discreteness results from the ranking given in the following tableau (where *Ca(*p*) represents *CATEG(*p*) for pitch value *p*, and *W(*d*) stands for *WARP(*d*) for distortion *d*). Most of the logically possible *CATEG constraints are undominated (shown in (i) in the figure); only two of the *CATEG constraints are sufficiently demoted to allow recognition of these tone categories (shown in (iii)). Large degrees of distortion are forbidden, since the demoted *CATEG constraints are outranked by *WARP constraints, most crucially in this case by *WARP(3) (see (ii)); thus [4] cannot be perceived as /1/. However, since, by Minimization of Distortion, the *WARP constraints disfavoring smaller distortions must be ranked below *WARP(3), distortion less than halfway between the two categories is tolerated; hence an input of [4] can be perceived as /5/. An input [3], falling directly between the two categories, would correctly be predicted to have no unambiguous categorization (i.e., /1/ and /5/ would be equally optimal percepts).

(5) Mature perception grammar for a system with two tones: /1/ (low) and /5/ (high)

	<i>i</i>			<i>ii</i>	<i>iii</i>	
[4]	*Ca(2)	*Ca(3)	*Ca(4)	*W(3)	*Ca(1)	*Ca(5)
/1/				*	*	
/2/	*					
/3/		*				
/4/			*			
☞ /5/						*

Categorical effects are also found in production. For example, the speakers of our

emerge, merely the number and phonetic details of the categories in any particular language. This seems reasonable to us; any list of the aspects of language requiring a nativist explanation should certainly include the ability to form categories.

two-tone language probably tend to pronounce the low tone with a pitch value close to [1] and the high tone around [5]. In Functional Phonology, this sort of phenomenon is also ultimately due to categorical perception, this time operating via the *REPLACE constraints. Like the *WARP family, *REPLACE constraints also conform to a functionally motivated principle, here designed to minimize categorization errors: “a production that gives rise to a less distant categorization error is preferred over one that leads to a more distant error” (Boersma 1998:177). This can be formalized as follows, with constraint ranking representing degree of preference and $|y-x|$ representing the difference between two values of some psychoacoustic feature.

(6) Minimization of Error

$$*\text{REPLACE}(\text{feature: } x, y_1) \gg * \text{REPLACE}(\text{feature: } x, y_2) \Leftrightarrow |y_1-x| > |y_2-x|$$

Speakers will thus tend to produce articulatory outputs that result in acoustic forms as close as possible to the underlying (i.e., perceptual) forms. However, physiologically it is too difficult to produce articulations that hit these targets precisely all the time. In the case of tone, such physiological motivations can be formalized with *GESTURE constraints like the following (where “tone” represents the complex articulations involved in producing a pitch, d measures the distance traveled by an articulator, and v measures its speed).

(7) *GESTURE (larynx: tone / distance = d , velocity = v)

The principles for the ranking of articulatorily motivated constraints in the production grammar all boil down to the extremely general one stated by Boersma (1998:149): “an articulation which requires more effort is disfavored,” represented schematically as follows.

(8) Minimization of Effort

$$*\text{EFFORT}(x) \gg * \text{EFFORT}(y) \Leftrightarrow x > y$$

This general constraint can be made explicit by grounding it in Newtonian mechanics. For example, changing a gestural articulation from one non-neutral position to another requires the movement of physical bodies, so (all other things being equal) longer and faster articulatory movements are more effortful (require more energy) than shorter and slower movements. These principles can be formalized as follows (see Boersma 1998:149-151 for a more thorough analysis describing further constraints and somewhat more complex notation).

- (9) a. *GESTURE(distance = d_1) >> *GESTURE(distance = d_2) $\Leftrightarrow d_1 > d_2$
 b. *GESTURE(velocity = v_1) >> *GESTURE(velocity = v_2) $\Leftrightarrow v_1 > v_2$

To see how these constraints and principles work, suppose in our two-tone system the target for production were /5.1.5/ (where “.” marks a syllable boundary). It is easier (for purely physiological reasons) to produce the low tone /1/ if its articulation is adjusted to be somewhat closer to that of the surrounding high tones. Given the universal rankings in (9), all tone languages will have a tendency to grammaticalize tonal coarticulation, though languages will vary in the amount of distortion from the articulatory target that is tolerated. In this theory, this variation is handled by the ranking of the *REPLACE constraints relative to the *GESTURE constraints. One possible ranking is shown below, where *R(p_1, p_2) represents a *REPLACE constraint forbidding the replacement of pitch p_1 for pitch p_2 , and *G($d=x$) represents a *GESTURE constraint forbidding movement between two tonal articulations x degrees apart, as measured for convenience with conventional pitch digits. Examination of the tableau in (10) shows that the *GESTURE and *REPLACE constraints compete with each other. As the candidates show greater amounts of coarticulation (from [5.1.5] to [5.2.5] to [5.5.5]), the ranking of the violated *GESTURE constraint drops (see (ii) and (iv)), but the ranking of the *REPLACE constraint rises (see (i) and (iii)). In this case, the intersection of the two competing rankings results in the selection of the optimal output [5.2.5], which shows some degree of coarticulation, but not total assimilation.

- (10) Mature production grammar for a system with two tones: /1/ (low) and /5/ (high)

	<i>i</i>	<i>ii</i>		<i>iii</i>	<i>iv</i>
/5.1.5/	*R(1,5)	*G(d=4)	*G(d=3)	*R(1, 2)	*G(d=0)
[5.1.5]		*			
☞ [5.2.5]			*	*	
[5.5.5]	*				*

2.2.2 Grammar, physics, and the lexicon in Functional Phonology

These Functional-Phonological analyses of the perception and production grammars may appear to be pure physics misleadingly couched in grammatical terms, but they are not. The above tableaux are intended to be representations of genuine phonological knowledge. As an interface system, it is crucial that phonological knowledge mirror physical reality as closely as possible, just as visual cognition and motor control (e.g.,

for reaching or walking) require mental representations and operations that are shaped by experience with the external physical world. The great advantage of the theory of Functional Phonology is that it can capture the knowledge of physics required for phonology in a way that is nevertheless also capable of handling categoricity and language variation (e.g., in the precise degree of tone coarticulation), all using a formalism familiar to most of the world's phonologists (i.e., Optimality Theory).

However, Functional Phonology has a fundamental weakness: it gives the impression that all of phonology is “natural”, which as Anderson (1981) and many others have pointed out, is not true. For example, velar softening in present-day English (e.g., *electric-electricity*) does not follow automatically from any obvious articulatory or perceptual principles. Yet this problem is hardly unique to Functional Phonology; unnatural patterns are a challenge to any phonological theory that attempts to be explanatory (e.g., Chomsky and Halle 1968 also have trouble with velar softening, since there is no way to go from /k/ to [s] in one step in their feature system). From a functionalist perspective, the key observation is the well-known correlation between the degree of unnaturalness of a pattern and its degree of lexicalization, as measured independently by the familiar diagnostics (e.g., the presence of lexical exceptions, interaction with morphology, structure preservation, and so forth; see lists in Hargus and Kaisse 1993:16 and elsewhere). Thus the more lexicalized a pattern, the less responsible the perception and production grammars are for describing its synchronic status. Nevertheless, if such patterns exist in the adult language, the learning algorithm described in Boersma and Hayes (2001) will automatically set the constraint rankings to encode them properly in the perception and production grammars, representing the knowledge as knowledge of lexical forms.⁷ In this paper we are primarily concerned with “natural” tone patterns, however, so the perception and production grammars formalized above will generally suffice.

3. Tone features

We now begin our analysis proper by applying the Functional Phonology model to two basic questions relating to tone features. In 3.1 we argue that the cross-linguistic limits on the number of level tones should be handled by constraints external to the representational mechanisms of the theory, rather than following directly from the

⁷ If there is indeed generative knowledge that goes beyond knowledge of perception and production, we suspect that it consists solely of lexical knowledge (e.g., analogy). Myers (2002a,b) present a general model of generative lexical knowledge in an OT framework; more specialized models include Benua (1997), Burzio (1997), Steriade (2000), and McCarthy (2001).

nature of tone features. In 3.2 we argue the same thing for the interactions between tone and other aspects of production, showing that a functional OT approach works better than one employing universal, articulatorily-based (i.e., laryngeal) tone features. In general, then, we argue that the behavior of tone follows from the mental representation of the physics of pitch and laryngeal articulation, not from abstract categorical features.

3.1 The number of level tones

It has always been standard in generative phonology to expect the formalism to account for cross-linguistic patterns of lexical contrast. For example, the apparent absence of any language showing more than a three-way contrast in vowel height led to the [high]/[low] system advocated by Chomsky and Halle (1968), while evidence for four-way and five-way contrasts later led Clements and Hume (1995) to challenge this feature system and replace it with a new one. Thus phonological theory should also be responsible for the fact that no tone language has been found that has more than five contrastive level tones (Maddieson 1978a). Yet as we argue in this section, the theory that handles this fact cannot be one that is purely formal, divorced of all influence by extra-grammatical factors.

Our analysis is adapted from Tsay (1994), who began with a more general observation from Maddieson (1978a): the more level tones in a system, the fewer the languages that exemplify that system (or more formally, there are fewer languages with $n+1$ contrasting level tones than n level tones).

(11) Calculated from data in Maddieson (1978a:366)⁸

Number of level tones	2	3	4	5	>5	Total size of sample
Number of languages	122	68	9	1	0	207

Most tone theories have not attempted to account for data like these, but instead make predictions for the maximum number of level tones that often bear little correspondence with the facts. Thus the maximum has been predicted variously to be three (Halle and Stevens 1971), four (Pulleyblank 1986, Bao 1990), five (Wang 1967, Sampson 1969, Woo 1969, Maddieson 1972), or nine (Hyman 1993, Duanmu 1990). Yip (1980), whose tone system was indeed designed to capture cross-linguistic limits on the number of level tones, does not fare very well either. She assumes that the true

⁸ Maddieson's sample also included seven languages with contrasting contour tones but no level tones. Tsay (1994:123-4) attempted to categorize these by the number of implied tone levels, but here we take a more conservative approach and include only actual level tones.

maximum is four, derived by combinations of her two binary features; the rare languages with five level tones she dismisses as misanalyzed or as requiring an extra, restricted feature. However, languages making five-way lexical contrasts in level tones are not dramatically more peculiar than any other tone systems, in spite of their relative rarity (a point also made by Odden 1995). Examples of such languages include Heimiao (Kwan 1966), Copala Trique (Hollenbach 1984), and Gaoba Dong (Shi, Shi, and Liao 1987). A minimal set of Heimiao words are shown below.

(12) From Kwan (1966), based on data collected by Fang-Kuei Li.

- a. [la¹¹] “candle; smooth, polished”
- b. [la²²] “to move away, to even”
- c. [la³³] “cave, den”
- d. [la⁴⁴] “a general classifier”
- e. [la⁵⁵] “short”

The more basic problem with categorical tone theories is that in the pattern in (11) there is no sharp cut-off as one might expect if the number of level tones were fixed by some categorical feature system (e.g., four-level systems common, five-level systems nonexistent or marked). Instead there is a gradual drop off, with four-level languages already rather rare. Given this pattern, the fact that six-level languages are (as yet) unattested should not really come as a surprise, nor would it be surprising if a more thorough search turned one up. Tsay (1994) thus concluded that the gradient drop-off must be accounted for by a gradient extra-grammatical factor, in particular memory, which would play an especially important role when the tone system was being learned. Each tone category would require positive evidence to learn, and learning would also be required to distinguish it from all of its competitors. This learnability hypothesis accounts for the fact that the drop-off is monotonic; the more tones there are to learn, the harder the system is to acquire, and the rarer it is.⁹

Since the formalism of Tsay (1994) was overly narrow (i.e., designed only to represent tone, not the acquisition of a tone system), it was not capable of doing more than simply stating this hypothesis in a general way. The model of Functional Phonology, however, can provide an account for the monotonic drop-off directly within its formalism, since the drop-off follows from Category Finiteness, the claim that only a finite number of *CATEG constraints can be dominated by *WARP constraints. As

⁹ As noted by an audience member at GLOW in Asia, this argument would be greatly strengthened by empirical evidence that multiple levels do in fact cause trouble for children acquiring a tone system. Evidence of this sort for the three-level tone system of Southern Min is described in Tsay (in preparation).

noted earlier, this claim itself is derived from learnability considerations: positive evidence about the necessity of a new phonemic category, and its prototypical form, is necessary in order to demote a *CATEG constraint from its original undominated position. Interestingly, as a reviewer pointed out, the framework of standard OT is already sufficiently functionalism-friendly to handle the drop-off pattern (as long as the key assumption of Tsay 1994 is also accepted, namely that tone features do not have inherent limits). One way is to use standard structure constraints that disallow categorical tone feature values, devices first proposed to handle phonemic inventories in Prince and Smolensky (1993, ch. 9). Like the *CATEG constraints of our analysis, these structure constraints are assumed to be undominated in the child's initial state (Tesar and Smolensky 1998), and thus it takes a greater amount of evidence for the child to demote the structure constraints needed to create a larger inventory. In either analysis, then, the monotonic drop-off emerges through language use (specifically acquisition), formalized with learning algorithms (Boersma and Hayes 2001 vs. Tesar and Smolensky 1998) that trace restricted paths through the space of all possible constraint rankings. Pure formalism would not be enough; factorial typology alone would falsely predict equal numbers of languages with large and small inventories, since unrestricted permutation would allow the structure (or *CATEG) constraints to be dominated and undominated equally often. By employing formal functionalism, however, the empirical limits on tone inventories can be captured without having to pass off the crucial explanations to a nonformalized part of the model.

3.2 The modality of tone features

Traditionally, features in generative theories tend to reflect articulation rather than perception (e.g., Halle 1983). In the case of tone, several attempts have been made to describe tone entirely with articulatory features, such as Halle and Stevens (1971), Ladefoged (1973), and Duanmu (1990). Bao (1990) presents a particularly ambitious analysis, proposing a feature geometry with the Laryngeal node dominating two separate nodes for each larynx-internal articulator: the Cricothyroid node for large adjustments in pitch (i.e., register), and the Vocalis node for finer adjustments of pitch within each register. Among other things, such proposals are meant to account for the affinity between voiceless consonants and adjacent high tones, and between voiced consonants and low tones, as has been revealed by sound change and synchronic patterns (see, e.g., Hombert, Ohala, and Ewen 1979).

As with limits on the number of level tones, however, we hold that the relation between tone and the larynx should emerge through the interaction of functionally motivated constraints, rather than reflecting the inherent nature of abstract tone

features.¹⁰ Particularly dramatic support for this position comes from the fact that tone can interact with non-laryngeal features as well. Such interactions are predicted because they are physiologically motivated, but since the motivation is not strong, they are correctly predicted to be rarer than interactions between tone and laryngeal features. In this section we show how Functional Phonology can formally capture these observations.

A prominent example of the relationship between tone and non-laryngeal features is found in the tone-vowel height interactions of Fuzhou (Wang 1969, Lehiste 1970, Wright 1983, Chan 1985, Tsay 1991, 1994, 1996, Tsay & Sawusch 1994, Jiang-King 1999). That is, the tone alternations of Fuzhou tone sandhi (too complex to discuss here) trigger vowel height changes as well: higher tones are associated with higher vowels. The following table summarizes all tone-triggered vowel alternations observed in Fuzhou.

(13) From Tsay (1991, 1994), based on data in Wang (1969), Yip (1980), and Chan (1985)

Tones	Vowel alternations								
LM, LML	æ	ai	au	ay	ei	ou	œy	ieu	uoi
H, HL, M	ɛ	ei	ou	oy	i	u	y	iu	ui

The key generalization is that when low tones (i.e., beginning with L) are replaced with higher tones (i.e., beginning with H or M), the vowels are raised one level of vowel height: low vowels ([æ] and [a]) become mid vowels ([ɛ], [e], or [o], assimilating the frontness and/or roundness of the offglide if present), while mid vowels ([e], [o], and [œ]) become high vowels ([i], [u], and [y], respectively). Analyses of this pattern that deny a direct tie between tone and vowel height, such as Wright (1983), Chan (1985), and Jiang-King (1999), focus instead on the apparent changes in prosody (e.g., diphthongs becoming monophthongs, or triphthongs becoming diphthongs). However, these alternative analyses neglect the facts that not all of the alternations change prosodic structure (e.g., [æ]-[ɛ], [ai]-[ei]), and that even when the number of segments does seem to change, this can always be understood as involving the merger of two high vowels (i.e., the raised mid vowels always produce a high vowel identical to the original preglide or offglide). Moreover, while interactions between tone and vowel height are relatively rare, they are not entirely unheard of; in addition to Fuzhou, they have also been reported in Cantonese (Hashimoto 1972), Hausa (Pilszczikowa-Chodak 1972, Newman 1975), Lahu (Matisoff 1973), and even Mandarin (Chao 1948, Tsay and Sawusch 1994).

¹⁰ Peng (1992) makes a somewhat similar point within the framework of Grounded Phonology (Archangeli and Pulleyblank 1994), a pre-OT theory of phonetic influences on phonology.

How can this interaction between tone and vowel height be accounted for? Clearly theories that claim that tone is handled by laryngeal features are insufficient, but it is equally unhelpful to propose that tone can have links in the feature geometry to vowel height as well. Adding such extra links would not only be difficult to represent geometrically, but it would miss the point that patterns such as Fuzhou's are intriguing precisely because they are relatively uncommon. Nevertheless, as with tone interactions with laryngeal properties like voicing, the explanation for the patterns like Fuzhou's also appears to lie in articulation. The articulation of the larynx is such that the raising of pitch height can be enhanced by moving the entire larynx upward with extrinsic muscles. This rotates the thyroid cartilage relative to the cricothyroid, thereby lengthening the vocal folds and raising pitch. This lifting of the larynx in turn raises the hyoid bone somewhat, to which the larynx is anchored at the top. The tongue root is also partly anchored to the hyoid bone, so raising this bone makes it easier to raise the tongue body as well. Thus raising pitch can reduce the energy required to produce a rise in vowel height. However, since the route from pitch raising to vowel raising is rather indirect, the physiological effect is rather small (Honda 1983, Honda, Hirai, and Dang 1994). These articulatory facts explain both why the Fuzhou pattern is able to exist, and why patterns like it are nevertheless relatively uncommon (see Tsay 1994 for further discussion).

The formalization of these observations thus requires a theory capable of encoding physical motivations in the grammar. Although the interaction of the gestures of distinct articulators is not examined by Boersma (1998), the insight to be captured is exactly parallel to the analysis of ease of articulation when only one articulator is involved. Just as it is easier to articulate a slightly raised low tone between two high tones than to articulate a very low tone in that context, it is likewise easier to articulate a high vowel when a high tone is simultaneously articulated. To express this in Functional Phonology, we shall adapt the articulatory coördination constraint mentioned briefly by Boersma (1998:156), as below. Note that this formalism is essentially an extension of the standard OT coöccurrence constraint (operating on categorical features) to phonetically detailed representations.

(14) *COORD(*gesture*₁, *gesture*₂):

The two gestures *gesture*₁ and *gesture*₂ are not coördinated.

Like *GESTURE constraints, *COORD constraints are subject to the ranking principles derived from Minimization of Effort. For example, in the case of tone, we can express the correlation between tone and voicing, and between tone and vowel height, with rankings of constraints like the following. (*Co stands for *COORD, /p/

stands for the articulatory gesture associated with any voiceless consonant, L, M, H represent three tone levels, and /æ/, /e/, /i/ represent three vowel heights.) The ranking in (15a) states that the lower the tone, the more disfavored it is to coördinate a voiceless consonant with that tone, while the ranking in (15b) states that the lower the height of a vowel, the more disfavored it is to coördinate a high tone with that vowel.

- (15) a. *Co(p,L) >> *Co(p,M) >> *Co(p,H)
 b. *Co(H,æ) >> *Co(H,e) >> *Co(H,i)

The *COORD constraints interact with *GESTURE constraints, which evaluate the effort involved in articulating gestures individually. Assuming that the articulation of extreme values (i.e., high or low vowels or tones) requires more effort than articulating mid values, we have the following simplified constraint rankings (using the same abbreviations as before).

- (16) a. {*G(H), *G(L)} >> *G(M)
 b. {*G(i), *G(æ)} >> *G(e)

To express the physiological facts that tone interacts directly with voicing, but not so directly with vowel height, we also need effort-based rankings between the *COORD constraints and *GESTURE constraints. The constraint ranking in (17a) below expresses the claim that it requires more effort to produce a mid tone with a voiceless consonant than it does to produce extreme tones (i.e., high or low) in general, since there are strong physiological forces to raise the mid tone in the context of a voiceless consonant. By contrast, the ranking in (17b) expresses the claim that producing a mid vowel with a high tone is not as difficult as producing extreme vowel heights (i.e., high or low) in general; since the articulation of tone has only a small effect on the articulation of vowel height, no coördination of them is particularly difficult. Both of these rankings should follow universally from Minimization of Effort, assuming that they manage to capture the physiological facts correctly in spite of being rather oversimplified.

- (17) a. *Co(p,M) >> {*G(H), *G(L)}
 b. {*G(i), *G(æ)} >> *Co(H,e)

We now apply these rankings to two schematic situations. The tableaux in (18) below evaluate candidate outputs varying respectively in tone-voicing interactions and in tone-vowel height interactions. They both involve *COORD and *GESTURE

constraints of parallel types, but there is a crucial difference in the ranking of the constraints highlighted under (ii) in the tableaux, corresponding to the different rankings given in (17). The result is that voicing will have a large effect on tone (i.e., favoring the output with an extreme tonal value, namely H), but tone will only have a small effect on vowel height (i.e., favoring the output with a mid vowel height).

(18) a. The large effect of voicing on tone

	<i>i</i>	<i>ii</i>			<i>iii</i>	<i>iv</i>
	*Co(p,L)	*Co(p,M)	*G(H)	*G(L)	*G(M)	*Co(p,H)
pa ^L	*			*		
pa ^M		*			*	
pa ^H			*			*

b. The small effect of tone on vowel height

	<i>i</i>	<i>ii</i>			<i>iii</i>	<i>iv</i>
	*Co(H,æ)	*G(æ)	*G(i)	*Co(H,e)	*G(e)	*Co(H,i)
æ ^H	*	*				
e ^H				*	*	
i ^H			*			*

Our analysis thus claims that tone-vowel height interactions are rarer than tone-voicing interactions because their effects are smaller. This means that they are less noticeable to children acquiring language, with the result that small deviations in vowel height are often “corrected” before they have any deep effect on the phonological grammar. In terms of Functional Phonology, the deviations are typically too small for *WARP constraints to force the emergence of a new categorization of vowels dependent on tonal context.

Once the constraints are established in the production grammar, however, new rankings can arise between the universally ranked *COORD and *GESTURE hierarchies and their competing *REPLACE hierarchies, and these may exaggerate the effect of tone on vowel height, possibly pushing the vowel into a different contrastive category, as in Fuzhou. Crucially, the new rankings in the production grammar must be driven by the perception grammar, since they result in productions that are more effortful, but which seem to be helpful for the identification of linguistically significant contrasts. The increase in effort is particularly clear in the case of Fuzhou, with its categorical shift

from one vowel phoneme to another (as demonstrated acoustically by Tsay and Sawusch 1994 and Tsay 1996), but even in the case of the relatively small, subphonemic tone-vowel height interactions of Mandarin, Tsay and Sawusch (1994) found that they were larger than what is predicted by the purely physiological model of Honda, Hirai, and Dang (1994). The extra cost in production is offset by benefits to the listener, since vowel height shifts aid in identification of contrastive tone categories by adding an extra co-varying perceptual feature. This perceptual explanation for exaggerated tone-vowel height interactions is supported by the lack of any measurable effect of pitch on vowel height in English, whose native listeners do not need to be able to identify lexical tone categories (Tsay and Sawusch 1994). The phonologization of such exaggerated interactions, then, begins in the perception grammar; Escudero and Boersma (2001) have shown that a very simple acquisition algorithm is capable of adjusting the rankings of constraints in a perception grammar so that it is sensitive to the precise degree of informativeness of different perceptual cues (e.g., the relative importance of F1 vs. duration in distinguishing /i/ and /ɪ/ in Scottish vs. Southern British English). These developments in the perception grammar then also cause adjustments in the production grammar via re-ranking of the *REPLACE constraints, which enforce identity between perceptual and articulatory targets.

In light of these considerations, note that in contrast to the implications of the tableaux in (18), Fuzhou mid vowels can appear with both low and high tones, depending on their role in the tone sandhi alternations. For example, the surface diphthong [ei] appears with low tones in juncture position, but with high tones in context position. This situation clearly cannot arise through the operation of the production grammar alone, but must have been forced by perceptually motivated restructurings that put two superficially similar surface forms into the same phonological category (i.e., the mid vowel of a surface juncture form [ei] and the partially raised low vowel of a target form /ai/ as produced in context position). Note the implicit past tense of the preceding clause: the separation of the production and perception grammars in Functional Phonology makes it impossible to represent the physiological motivation and perceptual categorization of a pattern within the same OT tableau, so a diachronic component is crucial in the explanation of patterns of this sort. Rather than a weakness, we consider this a strong argument in favor of this approach. A functionalist approach to phonology should not oblige one to deny the lexicalized nature of much of synchronic phonology (see also section 2.2.2 above and section 5 below). There is no contradiction in saying both that faithfulness constraints like *REPLACE play the crucial role in maintaining the vowel alternations in the production grammar of modern Fuzhou speakers, and that a complete explanation of these alternations requires reference to earlier stages in the development of the

production grammar when structure constraints like *COORD and *GESTURE held sway.¹¹

Space obviously does not permit exploration of these important and difficult issues here, but in any case we do not want them to obscure the key point of this section: in order to handle the interactions of tone with other phonological properties, Functional Phonology does not require that tone be represented with articulatory features of some particular sort. Instead, it provides a functionalist and constraint-based characterization of these interactions that is capable of predicting which interactions will be favored, and to what degree. This means that even relatively “marginal” phenomena like tone-vowel height interactions can find their proper place within phonological theory.

4. The autosegmental behavior of tone: articulation and contours

We now move beyond tone features themselves to their behavior in phonological processes. Specifically, we argue that tone spreading arises in the production grammar, thereby explaining an important restriction on the autosegmental behavior of contour tones. This restriction, simply put, is that contour tones never spread as wholes. It is perhaps unsurprising to discover this restriction in African languages, in which contours clearly behave like sequences of level tones (see review in Odden 1995), but from a purely autosegmental perspective it is unexpected for Asian languages, given the general tendency of such languages to treat tones as wholes (see review in Yip 1995). However, the observation has a straightforward explanation if the articulatory teleology of autosegmental spreading is recognized, as can be formalized in Functional Phonology.

Before getting to the explanation, we must first justify our claimed observation. Bao (1990), Yip (1989, 1995), and Chen (2000) have all argued that there are in fact examples in Sinitic languages showing that contours can spread, either as whole tones or as contours alone (i.e., tonal shape spreading independently of register). The languages that have been claimed to have such patterns are Zhenjiang, Wenzhou, Changzhi, Danyang, and Zhenhai. However, even after years of debate no consensus has emerged over the validity of the proffered examples. Thus Yip (1989, 1995) challenges Bao’s (1990) analyses of contour spread in Zhenjiang and Wenzhou, while Duanmu (1990, 1994) has carefully demolished all cases offered by Bao and/or Yip (including an earlier version of Yip 1995), i.e., Zhenjiang, Wenzhou, Changzhi, and

¹¹ It is perhaps possible to devise an analysis of Fuzhou that treats structure violations differently in different morphological contexts, something along the lines of comparative markedness (McCarthy 2002). To do this properly, however, we would have to get much deeper into Fuzhou tone sandhi than we have space for here, and we also suspect the solution may be merely technical, not an organic consequence of the functionalist approach.

Danyang. Even Chen (2000:76) notes that “[u]ncontroversial cases of contour spread, with or without involving the register, are indeed hard to find,” citing with apparent approval Maddieson’s (1978b) earlier claim that no rules of contour tone assimilation have ever been found. Doubt about putative cases of contour spreading is inevitable because the data are complex enough that alternative analyses are always available.

To provide some flavor of the inconclusive nature of the data, let us briefly examine Chen’s (2000) analysis of the Zhenhai tone sandhi pattern (to our knowledge not discussed by Bao, Duanmu, or Yip), in particular the subpattern found in compounds with a weak-strong stress pattern. Chen (2000), basing his analysis on data from Rose (1990), argues that this pattern demonstrates the spreading of contour independently of register. However, there are a number of problems with his argument, the most fundamental of which is that he has not described Rose’s data correctly (which is understandable, given the complexity of Rose’s tables and figures). The following table shows the basic tones and sandhi tones for the words listed by Chen (2000:68), with the pitch digits given in his figure (27) on the left, and the numbers given in Rose (1990) on the right (based on his Table 1 on p.3, Figure 2 on p.6, and Table 3 on p.8). Rose’s transcriptions correspond much more closely to the detailed acoustic phonetic data given in his paper, though even here quibbles can be made (e.g., the tone he transcribes as [441] should perhaps be transcribed as [431] or [421] to match the actual F0 contours in his graphs).

- (19) Tonal contours in pitch digits for disyllabic weak-strong words in Zhenhai; “.” marks syllable boundaries. Mismatches between Chen (2000) and Rose (1990) are highlighted.

Gloss		Chen (2000:68)	Rose (1990:3, 6, 8)
a. bedroom	basic sandhi	213 .441 11. 334	231 .441 11. 441
b. last year	basic sandhi	213.231 11.24	213.231 11.24
c. spring	basic sandhi	441.441 33.441	441.441 33.441
d. coal mine	basic sandhi	231. 231 11.441	231. 323 11.441

In Chen’s autosegmental analysis, the contour of the first syllable delinks, leaving it a level tone but maintaining its original register, and then the contour, as a unit, links to the second syllable, while maintaining the second syllable’s original register. However, the evidence is hardly compelling, even using his own (inaccurate) transcriptions. Chen’s transcriptions imply that the basic tone [213] has the same contour as both [334] and [24] (see (19a, b) above); differences in overall shape (e.g.,

the three-unit [213] and [334] vs. the two-unit [24]) and in slope (e.g., sharp rise in [213] and [24] vs. shallower rise in [334]) are left unexplained. Rose's more accurate transcriptions present similar problems for Chen. Merely adding phonetic implementation rules to supplement the autosegmental operations will not help. For example, [231] and [441] (in Rose's original transcriptions given above in (19a)) do not differ solely (or even primarily) in register, and in particular, the initial rise in [231] (absent in [441]) cannot be treated as phonologically irrelevant, since there is no obvious mechanical reason for its appearance here and its absence in [441]. The slope differences between [213] and [24] (in Rose's original transcriptions given above in (19b)) do not seem phonologically irrelevant either. It is true that initial dips like that seen in [213] are virtually universal with low tones (Maddieson 1978b), but the non-dipping low tone [11] seen in the same example implies that in this language, dipping or the lack thereof is under some degree of grammatical control.

A clue as to what may really be going on in Zhenhai can be seen in (19d), in which [441] appears as the sandhi form of the second syllable, even though this corresponds neither to the precise contour of the base-form first syllable [231] (even less so if it is more accurately transcribed as [431] or [421]), nor to the precise tone height of the base-form second syllable [323]. Instead, what it matches is the lexical tone [441], seen for example in the base form for (19a, c).¹² The appearance of [441] here suggests a "precompiled" aspect of Zhenhai tone sandhi (to use the term of Hayes 1990), in which one memorized tone may be substituted with another, rather than deriving one from the other through general rules or constraints. By itself, precompilation theory will not suffice for all of the Zhenhai alternations, most of which do not seem to be structure-preserving, but the appearance of an apparently lexical factor in this system may go some way towards explaining why a simple autosegmental spreading analysis of the sort attempted in Chen (2000) does not work.

Interestingly, lexicalization also seems to be a key factor in the other patterns that have been proposed as examples of contour spreading. Duanmu (1994) argues that the putative whole-tone spreading in Changzhi is more perspicuously analyzed as morphological tone reduplication, given that it affects precisely two morphemes (and

¹² Due to the confounding influence of adjacent segments, the acoustic phonetic data in Rose (1990) do not conclusively prove that what is transcribed as [441] is completely identical in words like (19a, c, d). For example, the F0 contours for words like (19a) vs. (19d) (which Rose identifies as 3+'1,2 vs. 3+'3,4, shown in part E of his Figure 3, p.15) are clearly affected by the fact that his sample (19a)-type words all have intervening voiceless consonants, while his sample (19d)-type words all have intervening voiced consonants; this raises the initial F0 contour for the second syllable in (19a)-type words and lowers it in (19d)-type words. Other than this perturbation, however, the contours appear to be the same. Clearly further phonetic work would be required to settle questions like this.

both happen to be function morphemes, which often have peculiar phonological behavior cross-linguistically). Similarly, Yip (1995:483) acknowledges that the “historical origins” of the key morpheme is relevant in the Danyang pattern, which suggests that it too is at least partly precompiled. Chen (1996:36) himself notes elsewhere that “apparently random tonal substitutions” are “quite common among Chinese dialects”. As noted earlier, such lexicalized and irregular phenomena should not be handled by the perception or production grammar *per se*. In any event, it has been far from conclusively demonstrated that contours can spread in Asian languages, and they certainly do not in other tone languages.

The observation established, we now turn to the explanation. Why can a tone contour not spread as a unit? Our answer has two parts. First, autosegmental spreading arises through the phonologization of coarticulation, i.e., the anticipation or perseveration of a gesture. This observation is of course not new; Kaisse (1992), Zsiga (1997), and Przewdziecki (2001) are among the many who have made the same point. For example, Kaisse (1992) points out that this hypothesis explains why the features whose autosegmental spreading is most widely attested (e.g., tone, nasality, vowel height) are associated with gestures that have been shown to be involved in long-distance coarticulatory effects due to the relative independence and/or slowness of their articulators. By contrast, the features that spread least readily, if at all (e.g., consonantality and sonorancy; see McCarthy 1988), are not associated with well-defined gestures of any sort.

The second part of our answer concerns the behavior of a contour tone undergoing coarticulation. If one extends the duration of the gesture(s) associated with a contour tone, the result is a contour with a longer duration relative to other gestures; it is impossible for repetitions of contour tones to arise through coarticulation. Putting the two points together, it is therefore impossible for the apparent autosegmental spreading of contour or whole contour tones to arise. What may easily arise is exactly what is widely attested: the splitting of a contour into a sequence of level tones, or the extension of a contour across a wider domain (e.g., in Shanghai; Duanmu 1997).¹³

To capture this analysis in terms of Functional Phonology, we merely have to build on the schematic production grammar presented in our introduction to this theory in section 2.2. The same *GESTURE constraints, used there to reduce the degree of change in tone articulations across syllables, can be used here to account for tone spreading patterns. The relevant constraint families and their universal rankings are repeated below.

¹³ Akinlabi and Liberman (2001) and Zhang (2002) precede us in applying phonetic insights to the distribution of contour tones, though not specifically to the problem of their (lack of) spreading.

- (20) a. *GESTURE(distance = d_1) >> *GESTURE(distance = d_2) $\Leftrightarrow d_1 > d_2$
 b. *GESTURE(velocity = v_1) >> *GESTURE(velocity = v_2) $\Leftrightarrow v_1 > v_2$

All else being equal, these constraint rankings favor gestures that move no distance at zero velocity, i.e., that are held in position.¹⁴ Disfavoring changes in gestural distance means that contour tones are disfavored relative to level tones, and that candidates containing different tones on contiguous syllables are disfavored relative to candidates with the same level tone on all syllables. Disfavoring rapid changes means that contour tones (where pitch changes within one syllable) are disfavored relative to having different level tones appearing on different syllables (where pitch changes across syllables).

The following tableau illustrates the effect of *GESTURE, free from the faithfulness force of *REPLACE (*G stands for *GESTURE, distance is given in differences in pitch digits, and velocity is given in number of syllables). The constraint under (i) is violated each time a change is made between [5] and [1]; (21g) [51.51] entails three such changes and so incurs three violations. The two constraints under (ii) are ranked in accordance to the principle in (20b): change within one syllable (velocity = 1 syllable) is worse than change across two syllables (velocity = 2 syllables). The candidate outputs are listed in order from most to least optimal; the focus here is not on picking a single optimal output, but showing how the constraints and their ranking impose a hierarchy on the candidates. Note that candidate (21g), with two identical contour tones, is the most disfavored.¹⁵

(21) Minimization of Effort for tonal sequences and contours.

		<i>i</i>	<i>ii</i>	
		*G(d=4)	*G(v=1σ)	*G(v=2σ)
a.	[5.5]			
b.	[5.1]	*		*
c.	[51.1]	*	*	
d.	[5.51]	*	*	
e.	[51.5]	**	*	*
f.	[51.15]	**	**	
g.	[51.51]	***	**	*

¹⁴ In Boersma (1998) they are countered by *HOLD constraints, not discussed here.

¹⁵ As a reviewer has pointed out, *G(d=4) is superfluous here since distance is a factor used in calculating velocity. However, it seems useful to be able to forbid movements of a given distance regardless of how slowly they are implemented. Otherwise gradual updrift across the entire pitch range (e.g., [1.2.3.4.5]) will falsely be predicted to be more common than shallow rising tones within a syllable (e.g., [35]).

Languages may show any of these tone patterns on the surface if forced by non-physiological constraints; even (21g) may appear if the two syllables each have a falling tone as their articulatory targets (as in Mandarin [pən⁵¹ tan⁵¹] “fool”), or if identity between aspects of the two syllables is required by morphological constraints (e.g., reduplication). *REPLACE constraints may also allow others among these outputs to be chosen as optimal, even if they do not match the input exactly. For example, if there is a high-ranking *REPLACE constraint requiring all tonal articulatory targets to be realized in actual production, a contour tone may be derived from a sequence of level tones if other constraints force the deletion of a syllable (e.g., /5.1/ → [51]).

Properly speaking, the above analysis is a model of coarticulation, albeit under grammatical control, not of spreading in the classic autosegmental sense. The major difference between the two is that autosegmental spreading produces output categories that are also used to make lexical contrasts. In the case of tone, this means that coarticulation of an underlying /51/ contour across two syllables will actually result in an articulation something more like [543.321], whereas the phenomenon typically analyzed as delinking and spreading would result in an articulation of [555.111]. From a purely production standpoint, [555.111] will always be disfavored relative to [543.321], since the former involves a much higher gesture velocity. To complete the analysis, then, the perception grammar must also play a role, converting [543] into /5/ and [321] into /1/ through categorical perception (as formalized in section 2.2). This process would increase the categoricity of the underlying forms involved in the production grammar. Then, through the faithfulness force of the *REPLACE constraints, the articulatory targets would themselves become much more categorical than “pure” coarticulation.

5. Conclusions and loose ends

In this paper we have tried to make the case that phonological theory should be simultaneously formal and functionalist: formal enough to make all claims explicit and empirically testable, yet functionalist enough to describe the interface nature of phonology accurately and completely. In our opinion, the most promising formal functional approach to phonology is functionalist OT, especially in versions like the Functional Phonology model of Boersma (1998), in which representations are phonetically detailed and constraints and their rankings are phonetically motivated, but which are designed to allow the emergence of categoricity (i.e., phonetic discreteness) and to permit cross-linguistic variation and lexicalization. In this paper we applied this model (modifying it where necessary) to tone, demonstrating how it accounts for limits on the number of contrastive level tones, interactions between tone and other aspects of

speech production, and the behavior of contour tones. Assumptions that are commonly made in the generative literature on tone (e.g., universal features that are categorical, autosegmental, and articulatorily defined) were shown to be unnecessary or incompatible with the facts.

As with any new approach, our model also raises questions. Most fundamental is the question that has appeared at various points in the discussion: what should be done with “unnatural” phonology? In this paper we implied that unnatural phonology can easily be distinguished from the sort of phonology that a functional model should be responsible for, but it does not always happen that phonological patterns divide so neatly into “natural” and “unnatural”. This is a very general problem, but one special case that deserves further discussion is the distinction between assimilation and dissimilation. While assimilation appears to have an articulatory teleology and hence appears to be “natural”, it has been argued by several independent researchers that dissimilation arises via perceptually driven restructurings of the lexicon, and hence may tend to be more “unnatural”. Thus Donegan and Stampe (1979) consider dissimilation to be a kind of fortition, motivated by perception rather than production since it helps to highlight details of the perceptual target. Boersma (1998) also argues that the Obligatory Contour Principle originates in the perception grammar, pointing out for instance that breaking up sequences of identical segments after morphological concatenation adds to articulatory cost but benefits the listener. Most relevantly, Ohala (1986) provides evidence that dissimilation can arise when listeners misperceive an intentional articulatory gesture as mere coarticulation and remove it from the underlying representation. This not only explains why dissimilation tends to involve features that can spread through coarticulation, but also, as noted by Kiparsky (1986) in his commentary on Ohala (1986) (see also Kiparsky 1995), it explains why dissimilations tend to be lexicalized (i.e., in his terminology, are lexical rather than postlexical rules).

The problem is that the analysis of dissimilations as perceptually motivated adjustments of the lexicon (hence “unnatural”) still faces some challenges that we have not yet fully resolved to our satisfaction. Take for example the following alternation seen in the Nikki dialect of Bariba (Welmers 1952), in which a word-final high tone (notated here as level 3) becomes a top tone (i.e., 4) before a low tone (i.e., 1).

(22) Welmers (1952:86-7), replacing tone diacritics with pitch digits.

- a. na³ ta³ su³ du² u² ra² “I planted yams”
- b. na⁴ ra¹ ta³ su⁴ du¹ u¹ re¹ “I plant yams”

In an earlier version of this paper we argued that a functional analysis of this apparent dissimilation was supported by the fact that the alternation is structure

preserving (Nikki Bariba has four contrastive tone levels), hence its behavior is consistent with that of a lexicalized pattern with an ultimately perceptual teleology. We neglected the observation of Welmers (1952:86, footnote 8) that the Kandi dialect of Bariba has precisely the same alternation, but there it is not structure preserving, since Kandi Bariba has no lexical top tone (i.e., no level 4).¹⁶ This does not mean that there is no functional analysis of Nikki and Kandi Bariba, just that we are not yet sure what it should be.

More generally, our model, like any functionalist model of phonology, seems to predict that phonetically natural phonology should be easier to process by adults than unnatural phonology if the degree of lexicalization is taken into account, but this prediction has never been carefully examined (though see Myers 2002a for some evidence that naturalness may not in fact play an important role in some kinds of phonological processing). As we noted in our earlier discussion, these problems are faced to some degree by all theories of phonology, and they are caused by an unstated assumption that the model presented in this paper shares with generative models in general: the descriptive theory (i.e., of speakers' knowledge about phonological patterns) is claimed to be entirely the same as the explanatory theory (i.e., of the ultimate origin of phonological patterns). To a large extent this simplifying assumption is sufficiently accurate to be useful, as we hope to have shown in this paper, but it is violated quite dramatically by lexicalized phonology, which may be part of speakers' synchronic knowledge in spite of having solely a diachronic explanation. A complete understanding of the interface system that is phonology must respect not only the tight constraints of physics, but also the far less constrained powers of human memory and pattern-learning. We leave these issues for future work, work that we hope other researchers interested in tone will share with us.

¹⁶ We are grateful to the theoretical criticisms from the reviewers that led us to reexamine the empirical details of Bariba more closely.

References

- Akinlabi, Akinbiya, and Mark Liberman. 2001. Tonal complexes and tonal alignment. *Proceedings of the North East Linguistics Society* 31, ed. by M. Kim and U. Strauss, 1-20. Amherst: GLSA.
- Anderson, Stephen R. 1981. Why phonology isn't 'Natural'. *Linguistic Inquiry* 12: 493-539.
- Archangeli, Diana, and Douglas Pulleyblank. 1994. *Grounded Phonology*. Cambridge: MIT Press.
- Bao, Zhiming. 1990. *On the Nature of Tone*. Cambridge: MIT dissertation.
- Bao, Zhiming. 1999. *The Structure of Tone*. Oxford: Oxford University Press.
- Benua, Laura. 1997. *Transderivational Identity: Phonological Relations between Words*. Amherst: University of Massachusetts dissertation.
- Boersma, Paul. 1998. *Functional Phonology*. The Hague: Holland Academic Graphics.
- Boersma, Paul. 1999. On the need for a separate perception grammar. University of Amsterdam Manuscript. Available at Rutgers Optimality Archive (roa.rutgers.edu).
- Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32:45-86.
- Burzio, Luigi. 1997. Surface constraints versus underlying representation. *Current Trends in Phonology: Models and Methods*, ed. by J. Durand and B. Laks, 97-122. Salford: University of Salford.
- Chan, Marjorie K. M. 1985. *Fuzhou Phonology*. Seattle: University of Washington dissertation.
- Chao, Yuen Ren. 1930. A system of tone letters. *Le Maître Phonétique* 45:24-27.
- Chao, Yuen Ren. 1948. *Mandarin Primer*. Cambridge: Harvard University Press.
- Chen, Matthew Y. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22:129-159.
- Chen, Matthew Y. 1996. Tone geometry—A Chinese perspective. *New Horizons in Chinese Linguistics*, ed. by C.-T. J. Huang and Y.-H. Audrey Li, 21-48. Dordrecht: Kluwer Academic Publishers.
- Chen, Matthew Y. 2000. *Tone Sandhi: Patterns across Chinese Dialects*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 1965. *Aspect of the Theory of Syntax*. Cambridge: MIT Press.
- Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper and Row. Reprinted 1990, Cambridge: MIT Press.
- Clements, G. N., and Elizabeth V. Hume. 1995. The internal organisation of speech sounds. *The Handbook of Phonological Theory*, ed. by J. A. Goldsmith, 245-306.

- Oxford: Blackwell Publishing.
- Donegan, Patricia J., and David Stampe. 1979. The study of natural phonology. *Current Approaches to Phonological Theory*, ed. by D. A. Dinnsen, 126-173. Bloomington: Indiana University Press.
- Duanmu, San. 1990. *A Formal Study of Syllable, Tone, Stress and Domain in Chinese Languages*. Cambridge: MIT dissertation.
- Duanmu, San. 1994. Against contour tone units. *Linguistic Inquiry* 25:555-608.
- Duanmu, San. 1997. Recursive constraint evaluation in Optimality Theory: Evidence from cyclic compounds in Shanghai. *Natural Language and Linguistic Theory* 15:465-605.
- Elman, J. L., and J. L. McClelland. 1988. Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically stored phonemes. *Journal of Memory and Language* 27:143-165.
- Escudero, Paola, and P. Boersma. 2001. Modelling the perceptual development of phonological contrasts with Optimality Theory and the Gradual Learning Algorithm. *Proceedings of the 25th Penn Linguistics Colloquium*. Also available at Rutgers Optimality Archive (roa.rutgers.edu).
- Flemming, Edward. 1995. *Auditory Features in Phonology*. Los Angeles: UCLA dissertation.
- Gussenhoven, Carlos, and René Kager. 2001. Introduction: Phonetics in phonology. *Phonology* 18:1-6.
- Hale, Mark, and Charles Reiss. 2000. Phonology as cognition. *Phonological Knowledge: Conceptual and Empirical Issues*, ed. by N. Burton-Roberts, P. Carr, and G. Docherty, 161-184. Oxford: Oxford University Press.
- Halle, Morris. 1983. On distinctive features and their articulatory implementation. *Natural Language and Linguistic Theory* 1:91-105.
- Halle, Morris, and Kenneth N. Stevens. 1971. A note on laryngeal features. *Quarterly Progress Report* 101:198-213. MIT Research Laboratory of Electronics.
- Hargus, Sharon, and Ellen M. Kaisse. (eds) 1993. *Phonetics and Phonology: Studies in Lexical Phonology*. San Diego: Academic Press.
- Hashimoto, O.-K. Yue. 1972. *Phonology of Cantonese*. Cambridge: Cambridge University Press.
- Hayes, Bruce. 1990. Precompiled phrasal phonology. *The Phonology-Syntax Connection*, ed. by S. Inkelas and D. Zec, 85-108. Chicago: University of Chicago Press.
- Hayes, Bruce. 1999. Phonetically driven phonology: The role of Optimality Theory and inductive grounding. *Functionalism and Formalism in Linguistics*, vol. 1: *General Papers*, ed. by M. Darnell, E. Moravcsik, M. Noonan, F. Newmeyer, and K. Wheatly, 243-285. Amsterdam: John Benjamins.

- Hollenbach, Barbara E. 1984. *The Phonology and Morphology of Tone and Laryngeals in Copala Trique*. Tuscon: University of Arizona dissertation.
- Hombert, Jean-Marie, John J. Ohala, and William G. Ewan. 1979. Phonetic explanations for the development of tones. *Language* 55:37-58.
- Honda, K. 1983. Relationship between pitch control and vowel articulation. *Vocal Fold Physiology*, ed. by D. M. Bless and J. H. Abbs, 286-299. San Diego: College-Hill Press.
- Honda, K., H. Hirai, and J. Dang. 1994. A physiological model of speech production and the implication of tongue-larynx interaction. *Proceedings of the Third International Conference on Spoken Language Processing (ICSLP'94)*, vol. 1 (S06-7): 175-178.
- Hyman, Larry M. 1993. Register tones and tonal geometry. *The Phonology of Tone: The Representation of Tonal Register*, ed. by H. van der Hulst and K. Snider, 75-108. Berlin: Mouton de Gruyter.
- Jiang-King, Ping. 1999. Tone-vowel interaction in Optimality Theory. *LINCOM Studies in Theoretical Linguistics*, vol. 16. Munich: LINCOM EUROPA.
- Kaisse, Ellen M. 1992. Can [consonantal] spread? *Language* 68:313-332.
- Keating, Patricia A. 1985. Universal phonetics and the organization of grammars. *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, ed. by V. A. Fromkin, 115-132. San Diego: Academic Press.
- Keating, Patricia A. 1990. Phonetic representation in a generative grammar. *Journal of Phonetics* 18:321-334.
- Kiparsky, Paul. 1986. Commentary on Ohala (1986). *Invariance and Variability in Speech Processes*, ed. by J. S. Perkell and D. H. Klatt, 400. Hillsdale: Lawrence Erlbaum Associates.
- Kiparsky, Paul. 1995. The phonological basis of sound change. *The Handbook of Phonological Theory*, ed. by John A. Goldsmith, 640-670. Oxford: Blackwell Publishing.
- Kirchner, Robert. 1997. Contrastiveness and faithfulness. *Phonology* 14:83-111.
- Kirchner, Robert. 1998. *An Effort-Based Approach to Consonant Lenition*. Los Angeles: UCLA dissertation.
- Kwan, J. C. 1966. *Phonology of a Black Miao Dialect*. Seattle: University of Washington MA thesis.
- Ladefoged, Peter. 1973. The features of larynx. *Journal of Phonetics* 1:73-83.
- Legendre, Géraldine, Jane Grimshaw, and Sten Vikner. (eds.) 2001. *Optimality-Theoretic Syntax*. Cambridge: MIT Press.
- Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge: MIT Press.
- Lieberman, Alvin M. 1982. On finding that speech is special. *American Psychologist* 37:148-167.

- Maddieson, Ian. 1972. Tone system typology and distinctive features. *Proceedings of the Seventh International Congress of Phonetic Sciences*, 958-961.
- Maddieson, Ian. 1978a. The frequency of tones. *Proceedings of the Berkeley Linguistics Society* 4:360-369.
- Maddieson, Ian. 1978b. Universals of tone. *Universals of Human Language*, vol. 2, ed. by J. H. Greenberg, 335-365. Palo Alto: Stanford University Press.
- Matisoff, James A. 1973. *The Grammar of Lahu*. Berkeley: University of California Press.
- McCarthy, John. 1988. Feature geometry and dependency. *Phonetica* 43:84-108.
- McCarthy, John. 2001. Optimal paradigms. University of Massachusetts Manuscript. Available at Rutgers Optimality Archive (roa.rutgers.edu).
- McCarthy, John. 2002. Comparative markedness. University of Massachusetts Manuscript. Available at Rutgers Optimality Archive (roa.rutgers.edu).
- Myers, James. 1997. Canadian Raising and the representation of gradient timing relations. *Studies in the Linguistic Sciences* 27:169-184.
- Myers, James. 2002a. An analogical approach to the Mandarin syllabary. *Journal of Chinese Phonology* 11:163-190.
- Myers, James. 2002b. Exemplar-driven analogy in Optimality Theory. *Analogical Modeling: An Exemplar-Based Approach to Language*, ed. by R. Skousen, D. Lonsdale and D. B. Parkinson, 265-300. Amsterdam: John Benjamins.
- Newman, Paul. 1975. The non-correlation of tone and vowel height in Hausa. *Studies in African Linguistics* 6:207-213.
- Odden, David. 1995. Tone: African languages. *The Handbook of Phonological Theory*, ed. by J. A. Goldsmith, 444-475. Oxford: Blackwell Publishing.
- Ohala, John J. 1986. Phonological evidence for top-down processing in speech production. *Invariance and Variability in Speech Processes*, ed. by J. S. Perkell and D. H. Klatt, 386-401. Hillsdale: Lawrence Erlbaum Associates.
- Peng, Long. 1992. *A Unified Theory of Tone-Voice*. Tuscon: University of Arizona dissertation.
- Pilszczikowa-Chodak, Nina. 1972. Tone-vowel height correlation and tone assignment in the patterns of verb and noun plurals in Hausa. *Studies in African Linguistics* 3:399-421.
- Port, Robert F. 1996. The discreteness of phonetic elements and formal linguistics: Response to A. Manaster Ramer. *Journal of Phonetics* 24:491-511.
- Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Rutgers University and University of Colorado Manuscript. Available at Rutgers Optimality Archive (roa.rutgers.edu).
- Przezdziecki, Marek. 2001. Vowel-to-vowel coarticulation in Yorùbá: The seeds of ATR

- vowel harmony. *WCCFL 19: Proceedings of the 19th West Coast Conference on Formal Linguistics*, ed. by R. Billerey and B. D. Lillehaugen, 385-398. Somerville: Cascadilla Press.
- Pulleyblank, Douglas G. 1986. *Tone in Lexical Phonology*. Dordrecht: D. Reidel Publishing.
- Rose, Philip. 1990. Acoustics and phonology of complex tone sandhi. *Phonetica* 47:1-35.
- Sampson, Geoffrey. 1969. A note on Wang's phonological features of tone. *International Journal of American Linguistics* 35:62-66.
- Sells, Peter. (ed.) 2001. *Formal and Empirical Issues in Optimality Theoretic Syntax*. Stanford: CSLI Publications.
- Shi, F., L. Shi, and R.-R. Liao. 1987. An experimental analysis of the five level tones of the Gaoba Dong language. *Journal of Chinese Linguistics* 15:335-361.
- Sproat, Richard, and Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21:291-311.
- Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. *Papers in Laboratory Phonology V: Acquisition and the Lexicon*, ed. by M. B. Broe and J. B. Pierrehumbert, 313-334. Cambridge: Cambridge University Press.
- Tesar, Bruce, and Paul Smolensky. 1998. The learnability of Optimality Theory. *Linguistic Inquiry* 29:229-268.
- Tsay, Jane. 1991. Combinatorial Specification: Evidence from Fuzhou vowels. *Proceedings of the International Symposium on Chinese Languages and Linguistics* 2, ed. by P. J-K. Li et al., 389-408. Taipei: Academia Sinica.
- Tsay, Jane. 1994. *Phonological Pitch*. Tuscon: University of Arizona dissertation.
- Tsay, Jane. 1996. Physical constraints in the phonology of Fuzhou. *Proceedings of the Seventh North American Conference on Chinese Linguistics and the Fifth International Conference on Chinese Linguistics*, ed. by T.-F. Cheng, Y. Li, and H. Zhang, 344-361. Los Angeles: GSIL, University of Southern California.
- Tsay, Jane. (In preparation) Neural network modeling of tone acquisition in Taiwanese. National Chung Cheng University Manuscript.
- Tsay, Jane, and James R. Sawusch. 1994. The interaction of F0 and vowel quality in tonal and non-tonal languages. *Journal of the Acoustical Society of America* 95.5, pt. 2, 2922.
- Wang, T.-C. 1969. *Fuzhou Yuyin Yanjiu [Phonetic Research on the Foochow Dialect]*. Taipei: Sun Yat-Sen Cultural Foundation.
- Wang, William S.-Y. 1967. Phonological features of tone. *International Journal of American Linguistics* 33:93-105.
- Welmers, William E. 1952. Notes on the structure of Bariba. *Language* 28:82-103.
- Woo, Nancy. 1969. *Prosody and Phonology*. Cambridge: MIT dissertation.

- Wright, Martha. 1983. *A Metrical Approach to Tone Sandhi in Chinese Dialects*. Amherst: University of Massachusetts dissertation.
- Yip, Moira. 1980. *The Tonal Phonology of Chinese*. Cambridge: MIT dissertation. Reprinted by Garland 1990 as *The Tonal Phonology in Chinese*.
- Yip, Moira. 1989. Contour tones. *Phonology* 6:149-174.
- Yip, Moira. 1995. Tone in East Asian languages. *The Handbook of Phonological Theory*, ed. by J. A. Goldsmith, 476-494. Oxford: Blackwell Publishing.
- Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.
- Zhang, Jie. 2002. *The Effects of Duration and Sonority on Contour Tone Distribution: Typological Survey and Formal Analysis*. London: Routledge.
- Zsiga, Elizabeth C. 1997. Features, gestures, and Igbo vowels: An approach to the phonology-phonetics interface. *Language* 73:227-274.

[Received 26 April 2002; revised 31 December 2002; accepted 6 January 2003]

James Myers
Graduate Institute of Linguistics
National Chung Cheng University
Min-hsiung, Chiayi 621, Taiwan
Ingmyers@ccu.edu.tw

Jane Tsay
Graduate Institute of Linguistics
National Chung Cheng University
Min-hsiung, Chiayi 621, Taiwan
Ingtsay@ccu.edu.tw

形式功能的聲調模式

麥 傑 蔡素娟

國立中正大學

對語言的研究，形式主義與功能主義長久以來都被認為是互不相容的。然而近年來，一方面形式學家逐漸意識到語言使用對語言結構的影響，另一方面功能學家則提出較明確的運作模式以測試這種影響，這個互不相容的情勢已經開始改變。本文所提出的聲調模式是形式的，也是功能的。這個模式是形式的，因為它的目標在提供一個十分精確的音韻知識的理論；它也是功能的，因為模式中的運作成分的形式都是來自說、聽、習得等音韻知識在語言使用時的實際情形。本文最主要的是採用 Tsay (1994) 對聲調現象的洞察與 Boersma (1998) 的功能優選理論。這個新的聲調理論可以解釋聲調中各種重要的現象，例如：語言所用到的平調的數目上限；聲調與其他音韻特徵的互動；聲調傳播等。語料主要來自漢藏語系的聲調語言。

關鍵詞：音韻學，聲調，功能主義，優選理論