# Contrast and Complexity in Chinese Tonal Systems

by

Ross Krekoski

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Graduate Department of Linguistics
University of Toronto

# Contrast and Complexity in Chinese Tonal Systems

Ross Krekoski

Doctor of Philosophy

Graduate Department of Linguistics
University of Toronto

2017

## Abstract

Lexical tone has been investigated from a diverse variety of perspectives, with significant disagreement as to whether a feature-based analysis can account for the synchronic and diachronic patterning of tonal systems. This thesis investigates the tonal systems of 13 Chinese languages from the Yue and Mandarin groups, and presents a feature-based model which accounts for their respective inventories and patterns of phonological alternations, and traces their respective diachronic trajectories back to the feature specification of a reconstructed Middle Chinese tonal inventory.

The analysis here employs hierarchically arranged contrastive features. Whereas analyses postulating a universal set of phonetic features tend to incur a number of descriptive and theoretical problems when tasked with the description of tonal systems, utilizing language-specific features grounded in tonal alternations and phonetics sheds significant light on synchronic and diachronic patterning.

The reconstruction of a common Middle Chinese proto-language enables the description of a precise phylogenetic tree as well as the postulation of a number of contrastive reanalyses. It is found that both reanalysis and the specific phonetic content of tonal inventories are accounted for by postulating a principle of complexity minimization: the cognitive apparatus chooses representations which minimize the algorithmic entropy or Kolmogorov complexity of phonologies and maximize the reuse of structure. At each point in the diachronic trajectory of a language, speakers choose the phonological representation with the lowest total complexity. Phonetic naturalness can also be formalized as the minimization of complexity, and a general tendency towards an inverse relationship between the naturalness of tonal inventories and the amount of activity in a given language is described. The cognitive preference for the minimization of complexity represents a parsimonious solution to the problem of induction which avoids the stipulation of hard-coded universals pre-specified in the cognitive apparatus.

# Acknowledgments

I cannot thank my supervisor, Elan Dresher, enough. One could say that this thesis would not have been possible without his foundational work on contrast providing the inspiration for this project, his incredibly vast knowledge of the intellectual history of phonology informing said inspiration, and his intellectual generosity in giving space for the entertainment of unconventional hypotheses (many of which were, rightfully, abandoned along the way). One could say this, but it would be an understatement. I am grateful most of all for all of our fun, informative, and very wide-ranging conversations on ‘The Big Picture’: I have learned so much, and being given the opportunity to even conceptually approach this project from so many directions was a true privilege.

I similarly owe a tremendous debt to the other members of my committee: to Keren Rice and Yoonjung Kang for their absolutely tireless energy, endless feedback, helpful insights and kind support and encouragement from the very beginning of this project; and to Peter Jurgec, Michael Barrie, and Zhang Jie for reading and providing feedback, and serving on the examination committee. I am truly thankful for the excellent contributions and stimulating and very constructive discussion at the defense.

I am very grateful for the opportunity to have been a graduate student in the Department of Linguistics at the University of Toronto. This department is truly special. Amongst faculty, staff, and my fellow students over the years there are too many people who’ve provided cheer, support, laughs, and awesome conversation to thank individually. My years spent here have been incredible.

Finally, I must convey my deep gratitude to my family. To mom and dad for their ceaseless support and encouragement and contagious confidence. To Erin, Caitlin, Brianna, and Taylor for always being there to listen, chat, laugh, remind me that there is a world outside of research, and sometimes even humor my enthusiasm for the obscure by feigning interest in it. And, of course to my dearest wife, Lynda, and daughter, Caerys for walking (and, I suppose, crawling) this road with me and being the very brightest lights in my life. There are no adequate words.

# Table of Contents

# List of Tables

# List of Figures

# 1

# Introduction

This thesis describes the tonal systems of 13 Chinese languages. It has several intersecting goals. Descriptively, it seeks to provide an account of tonal inventories and processes in each of these languages, both individually and cross-linguistically, as well as to account for their development from Middle Chinese. This study also has a number of theoretical goals. It seeks to explore the degree to which a theory of contrastive features can account for the synchronic and diachronic behavior of tonal systems. It explores the role of phonetic substance in a theory of emergent features, particularly as it relates to diachronic change. It highlights and explores the interrelationships between phonetic substance, naturalness, and phonological activity. It also seeks to incorporate notions of formal complexity and cognitive simplicity into an overarching account.

The study of tone has been approached from a wide variety of perspectives, and there is disagreement as to whether tonal systems can be explained by a theoretical model involving features. While early research (Gruber 1964; Wang 1967) assumed features, Goldsmith's (1976) highly influential research on African tonal systems culminating in the autosegmental approach to tone set the stage for much subsequent inquiry to treat tones as complex composites of flat primitives, rather than unitary entities. Chinese languages, however, tend to exhibit both larger inventories and much more complex sets of tonal processes than African tonal languages, and may not be suited to an autosegmental approach. This is explored in Chapter 2, which also provides a brief outline of tonal systems in Chinese languages. I suggest that the framework associated with the contrastive hierarchy (Hall 2007; Dresher 2009) provides a promising foundation for describing tone in these languages.

In Chapters 3 and 4 I analyze a number of tonal systems in Mandarin and Yue languages. We find not only that a theory of contrastive features can account for tone in these languages, but that it offers significant insight into their synchronic and diachronic behavior, and we are able to account for the modern structure of each language analyzed by postulating a small number of sequential contrastive changes to a common ancestor. Phonological activity is observed to be closely associated with principles of phonetic naturalness: we observe here that there is an inverse relationship between the degree to which a given inventory is phonetically natural and the number of tonal alternations exhibited by a given language. That is, tonal inventories tend to be the most natural in languages with the fewest number of tonal alternations. We also make the observation that structural reanalyses tend to be conditioned by the principled maximization of phonetic naturalness.

Noting the large role that phonetic naturalness appears to play in the synchronic and diachronic behavior of the languages investigated here, I formalize the tendency towards maximizing naturalness, mediated by a number of foundational constraints, in Chapter 5. I discuss findings which suggest that the cognitive apparatus attempts to minimize complexity, or algorithmic entropy in its representations, and that this preference for simplicity cuts across numerous ostensibly disparate cognitive domains. I argue that phonetically natural representations can be conceptualized as representations which are at complexity minima, and that the preference for complexity minimization conditions many specific aspects of diachronic change.

The general cognitive preference for complexity minimization represents a candidate solution to the problem of induction, the observation that there are an infinite number of patterns to which a given set of data conform. The observation that any of an infinite number of formal systems can generate a given set of data represents a theoretical problem of importance to studies of acquisition, leading researchers (e.g. Chomsky and Halle 1968) to stipulate universal conventions pre-encoded into the cognitive apparatus as a means to place bounds on this otherwise infinite system space. The present study both presents a more parsimonious alternative, being derived from domain-general cognitive preferences and thus not requiring the stipulation of excess theoretical entities, and also does so in accordance with a broader literature of psychological observations and cognitive and information-theoretic prediction.

The scope of this study comprises an analysis of the tonal systems of 13 Mandarin and Yue languages. The Mandarin and Yue language groups were chosen for two reasons: i) they are, by a significant margin, the most well described and understood Chinese language groups in the literature; and, ii) the two groups, while both originating from a common ancestor in Middle Chinese, exhibit interesting typological differences pertaining to tone: Mandarin languages possess an unusual abundance of phonological processes which are somewhat anomalously near-absent in the Yue languages, a difference which proves critical to the present work. Languages within each group were similarly chosen which have a high amount of descriptive scholarship and which possess unique and diverse inventories and/or processes. Aside from a few notable exceptions, Chinese languages tend to be under-described, and, descriptions which do exist often disagree on the precise details of tonal inventories and processes. Theoretical explanation tends towards being even more diverse than simple description, highlighting the fact that the debate on how to treat lexical tone is simply not settled to any satisfactory extent in general, and this is even more the case for Chinese tonal systems. The extent of this is captured by Matthew Chen, who states in his excellent and influential treatise, *Tone Sandhi,* that:

> there remains a vast assortment of tonal alternations that defy classification and description let alone explanation. As one examines one Chinese dialect after another, one is left with the baffling impression of random and arbitrary substitution of one tone with another without any apparent articulatory, perceptual, or functional motivation (Chen 2000:81-82)

While I leave the question of motivation to future inquiry, this study finds that the distribution of alternations and the structures of inventories in the languages investigated, far from being baffling, random, or arbitrary, are completely predictable from a few basic principles. The present study incorporates all 13 languages into a unified framework which sheds new light both on their descriptive specifics as well as on their respective structural similarities, differences, and diachronic changes from Middle Chinese.

# 2

# Tone, Features, and Contrast

## 2.1 Overview

This chapter will provide some background required for a discussion of tonal systems in Chinese languages. I begin in §2.2 with a brief discussion of the tonal categories and historical scholarship of Middle Chinese, from which all languages discussed here are descendent. This is followed by a discussion of the representation of tone in various contemporary theoretical frameworks and an outline of some of the motivations for theoretical assumptions of the present work in §2.3.

## 2.2 Chinese language groups

Modern Chinese languages are generally classified into several major language groups. Numerous surveys regarding group characteristics, genetic classifications and distribution have been done (Egerod 1967; Matisoff 2003; Norman 1988; You 1992; Yuan 1960). The *Language Atlas of China* (Wurm et al 1987) is likely the most complete culmination of these efforts, and classifies Chinese languages into ten distinct groups, as in Table 1 below.

| Group | Speakers (in millions) | Location | Representative languages |
|---|---|---|---|
| Mandarin | 662.2 | north of Yangzi rivers, southwest provinces, worldwide in diaspora | Beijing, Tianjin |
| Jin | 45.7 | Shanxi, west Hebei | Pingyao, Changzhi |
| Wu | 69.8 | south Jiangsu, Zhejiang, south-east Anhui | Shanghai, Suzhou, Danyang |
| Hui | 3.1 | south-east Anhui, west Zhejiang | Tunxi |
| Gan | 31.3 | Jiangxi, east Hunan | Nanchang |
| Xiang | 30.9 | Hunan | Changsha |
| Min | 55.1 | Fujian, Taiwan, east Guangdong, Hainan, southeast Asia in diaspora | Fuzhou, Xiamen, Taiwanese |
| Yue | 40.2 | Guangdong, east Guangxi, south-east Asia, worldwide in diaspora | Cantonese, Taishan |
| Pinghua | 2.0 | south Guangxi | Nanning |
| Hakka | 35.0 | south Jiangxi, west Fujian, East Guangdong, Taiwan | Meixian, Changting |

Table 1: Chinese language groups (adapted from Chen 2000; Wurm et al 1987).

Chen (2000) writes that we can make meaningful typological generalizations, specifically in terms of tone, not just regarding individual languages, but regarding groups as well. For instance, the Southern dialects (for example the Yue group) typically have larger inventories than the Mandarin group (Chen 2000; C. Cheng 1973, 1991). The Jin, Wu, Min, Hakka and Mandarin languages as a class tend to exhibit “highly complex tone sandhi, while Xiang, Gan, and especially Yue show only limited (if any) tonal alternations” (Chen 2000;2). Further, sandhi processes often tend to take different forms in different language groups: “tone deletion and tone spread, widely attested in Wu, are all but unknown among Mandarin and Min” (Chen 2000:2).[1]

All of the languages discussed in the present study descend from Middle Chinese, spoken around AD 200 to 900 (Chen 2000:5). The traditional Chinese philological tradition, dating at least as far back as the 7th century, treated tones in Chinese languages as unitary entities. In the *Qieyun*, a pronunciation dictionary (AD 601), tone was described in terms of register: *yin* and *yang*, for high and low registers respectively, and general tonal shape or contour: *ping*, 'even' or 'level'; *shang*, 'rising'; *qu*, 'departing'; and *ru*, 'entering'. Each of these tonal shape descriptors were understood to describe a tonal category, referred to contemporaneously as Middle Chinese tones I, II, III, and IV, respectively, by modern scholars (Bao 1999; Chang 1975; Pulleyblank 1978, 1984).[2] In traditional and much contemporary Chinese scholarship, syllables in Chinese languages

---

1 Nantong, a Mandarin language which we discuss in §3.2.6 below, is a notable exception, exhibiting both tone deletion and spread, likely due to heavy influence from Wu languages, which it borders.

2 While traditional scholarship treated each as a separate category, modern scholarship treats tones I, II, and III as contrastive and IV as allotonic. Middle Chinese thus possessed three contrastive tones.

are classified into ‘checked’ and ‘smooth’ syllables, with the former ending in an occlusive coda. Middle Chinese tone IV was restricted to checked syllables. Each individual tone was described with at least a shape descriptor and, if necessary, a register descriptor (as in (1) below). Wang writes that although Middle Chinese is traditionally regarded as having four distinct tones (I argue in §3.2.7 that it initially possessed only three contrastive tones), it likely had eight pitch shapes, with four low pitch variants appearing on syllables with voiced initials, and the remaining four high variants on syllables with voiceless initials (1967:95).[3] Although the shape or contour descriptor, at one point, most likely gave descriptive phonetic information regarding tonal shape, the descriptors are now generally used to refer to specific tonal categories: when applied to modern Chinese languages, the descriptors refer to the modern correlate[4] of a given tone in Middle Chinese, regardless of phonetic shape. For example, the modern correlates of the *yin ping* (high-even) tone, (Middle Chinese Ia) is realized as high-flat in Beijing Mandarin, but low rising in Pingyao, and low falling in Tianjin. Although these categorical descriptors give us very little useful information regarding the actual current phonetic shape of a given tone, they do illustrate two points implicit in the Chinese linguistic tradition. Firstly, tones were understood to have individual identities, rather than being composites (as in, for example Leben 1978 or Sagey 1986, discussed below). Secondly, it was understood that certain phonetically identifiable dimensions served to contrastively differentiate these individual tones: the very fact that, for example, a high-low

---

3 Subsequent devoicing of initials led to the later introduction of additional tonal contrasts in some language groups.

4 Of course this is not a clear cut designation, as there have been multiple tone mergers, splits, and redistributions since Middle Chinese.

register descriptor existed at this time suggests that register was understood to be of some consequence in tone differentiation.

## 2.3 Representations of Tone

Chao (1930) first introduced the numerical notational system for describing the phonetic shape of tone. In this system sequences of up to three digits with values ranging from 1-5 are used to denote the overall height and shape of a tone. A high falling tone thus can be described with 53, or 41, a low flat tone could be described with 22, and a low concave tone could be described with 214, among other numerous possibilities. This is illustrated in the example below (adapted from Bao 1999:12) from Songjiang, spoken in Shanghai, which also includes the traditional Chinese descriptors in the first column and categories in the second. Interestingly, both modern correlates of Middle Chinese tone I (that is, Ia and Ib), which traditionally was described as 'flat' or 'even' are falling, and both correlates of tone II (described traditionally as falling) are flat.[5] In this particular language, low register predictively co-occurs with consonant initial voicing:

---

[5] The tendency for and degree to which languages maintain a given structure of contrasts, even in the face of significant tonal drift, will be revisited at length throughout this work.

(1) Tonal inventory of Songjiang

| Middle Chinese descriptor | Category | Tone | Songjiang | English |
|---|---|---|---|---|
| *yin ping* (high-flat) | Ia | 53 | *ti* | 'low' |
| *yin shang* (high-falling) | IIa | 44 | *ti* | 'bottom' |
| *yin qu* (high-rising) | IIIa | 35 | *ti* | 'emperor' |
| *yin ru* (high-short) | IVa | 5 | *paʔ* | 'hundred' |
| | | | | |
| *yang ping* (low-flat) | Ib | 31 | *di* | 'lift' |
| *yang shang* (low-falling) | IIb | 22 | *di* | 'brother' |
| *yang qu* (low-rising) | IIIb | 13 | *di* | 'field' |
| *yang ru* (low-short) | IVb | 3 | *baʔ* | 'white' |

Bao (1999) points out that this system overgenerates possibilities: a numerical scale with five positions and three digits allows for 125 possible combinations, and no language possesses this many phonologically contrastive tones (Cheng (1973) extends this system to a 7 position height scale). In addition, while languages with four contrastive height distinctions and languages with five phonetically distinct pitch levels both exist, Bao argues (1999:12) that those which possess five phonetically *and* phonologically distinct tone levels (for example, those possessing five contrastive flat tones differing only by tone height) have yet to be documented and that the system also fails to give any phonologically relevant information. It should be noted that, although it may superficially appear to represent complex tones as composites of sequential simplex tones, Chao's system, as originally conceived views tones as unitary entities.

All this aside however, this system will be adopted for the present work. Chao's numerical notational system will be utilized with forward slashes to indicate tonemes (e.g. /213/ indicates the low concave tone in Beijing Mandarin). This preserves phonetic information regarding acoustic properties of tone while underscoring a conceptualization of tones as unitary entities.

### 2.3.1 Tone and Features

Feature-based descriptions of tone have traditionally been fairly intimately associated with, and often motivated by, research regarding observations of universals of tonal systems. Regarding these, Maddieson (1978b) and Yip (1980) and later Bao (1990, 1999) make the following observations:

(2) Universal observations of tonal languages:

a. Tonal languages do not exhibit more than four level, and four contoured tones.

b. If contoured tones are present, there will be no more than two rising and two falling tones.

c. If complex[6] tones are present, they supplant simple contoured tones such that no more than four contoured and eight total tones are present in a language.

Wang's (1967) system is generally regarded as the first attempt of many to incorporate observations of the supposed universal boundary conditions (whether innately or functionally

---

6 Complex tones are concave or convex tones, where a single tone rises and then falls, or falls and later rises (e.g. the low concave tone in Beijing Mandarin, /214/).

specified) of tonal systems into a feature-based analysis.[7] He states that any tonal system needs to account for the potential existence of a five-way height distinction,[8] two rising tones, two falling tones, two convex and two concave tones. He proposes the following feature-based specifications and tones:

**Tone**

| | ˥ | ˩ | ˦ | ˨ | ˧ | ˧˥ | ˩˧ | ˥˧ | ˧˩ | ˥˧˥ | ˧˩˧ | ˧˥˧ | ˩˧˩ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Feature** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| contour | − | − | − | − | − | + | + | + | + | + | + | + | + |
| high | + | − | + | − | − | + | − | + | − | + | − | + | − |
| central | − | − | + | + | + | − | − | − | − | − | − | − | − |
| mid | − | − | − | − | + | − | − | − | − | − | − | − | − |
| rising | − | − | − | − | − | + | + | − | − | + | + | + | + |
| falling | − | − | − | − | − | − | − | + | + | + | + | + | + |
| convex | − | − | − | − | − | − | − | − | − | − | − | + | + |

Table 2: Wang's (1967) tone feature specification.

This set of specifications is not meant to imply that any given language must be able to generate all 13 tones in the above table and thus necessarily specify all seven possible features, but simply that this set of features is sufficient to descriptively account for all attested tones using

[7] Anderson (1978) describes an unpublished work by Gruber (1964) as the first description of tone which utilized features. Gruber distinguishes between pitches above or below a mean value with the feature [high]. Each of these ranges are further differentiated by means of a [high2], or 'secondary high' feature, and the placement of contoured, that is rising, falling, concave or convex tones into one of the four resultant categories was determined by whether 'most' of the contour was above or below a median range. Anderson (1978) argues that this analysis is somewhat inadequate and not as phonetically 'realist' as subsequent inquiries.

[8] It is unclear that a five-way height distinction is necessary. Bao (1999) argues that four is all that is necessary. Maddieson (1978) argues five is the maximum number of contrastive distinctions that may be made, but states that a number of what are apparent five-tone languages have at least one tone which serves as a predictable variant of another. Neither outcome bears directly upon the present analysis.

some combination of the above. Wang proposes that both these features and tones are universal, and while this entails a certain degree of redundancy in the system, he argues that a set of language-independent redundancy conventions minimizes the overall complexity of the paradigm. We can thus read Table 2 as partially derivative of a set of features and a set of redundancy conventions. While the theory attempted to treat tone via the exact same mechanisms as segmental features, it eventually ran into issues in the precise determination of which tones and features are utilized by which languages. Wang argues that:

> Languages with three tones would characteristically have tones 1, 2, and 5… or perhaps 6, 8, and 5; we would hardly expect them to have, say, tones 1, 3 and 6 or 1, 10 and 11. Languages with four tones may have tone 1, 2, 6, and 8 or 1, 2, 8, and 13; but hardly ever 1, 2, 3, and 4 or 6, 7, 8, and 9. (1967: 103)

This determination, according to Wang, is made with an accompanying set of universal markedness conventions, following then-contemporary work by Chomsky and Halle (culminating in the *Sound Pattern of English* in 1968), such that each feature in a given tonal context is given a default value: +, -, *u*, or *m*, with all values save for *unmarked* adding to the total complexity of a system, and the system attempting to minimize complexity in all cases. While the system worked quite well at accounting for the inventories and processes for the small set of languages investigated, it was not well motivated phonetically, formally complex and contrived (Clements et al. 2010), and subsequent research found the work to have "no solid crosslinguistic basis and it was quickly and widely abandoned" (Clements et al. 2010; see also critiques by Stahlke 1977 and Chen 2000).

Barrie's (2007) work on tone in Chinese languages also draws heavily on universal observations regarding tone. He proposes that three features can capture the above-mentioned universals on tone. According to Barrie, the register feature [upper] or [lower] divides the entire pitch range into registers, the pitch features [hi] or [lo] serves to further divide each half of the register in two, yielding a four way height distinction, and the [contour] feature determines whether a given tone is flat or contoured. Barrie's argument that pitch in Chinese languages can be represented by these three features is based on a proposal by Yip (2001) that contour tones are only specified for pitch at the left edge. That is, while the pitch value at the left edge is necessary to make distinctions between tones, the final pitch value is irrelevant, and aside from the initial pitch value, only directionality or shape serves to contrastively differentiate tones. Barrie argues thus that in the languages examined, the pitch division serves to predict whether a given tone is rising or falling, as in (3) below (adapted from Barrie 2007:349-350):

(3) Tones in Guangzhou Cantonese

| | | | | | |
|---|---|---|---|---|---|
| a. | high-falling | /53/ | *saan* | 'to close' | [upper][hi][+contour] |
| b. | high-level | /55/ | *gam* | 'gold' | [upper][hi][-contour] |
| c. | high-rising | /35/ | *hou* | 'good' | [upper][lo][+contour] |
| d. | mid-level | /33/ | *sei* | 'four' | [upper][lo][-contour] |
| e. | extra-low-level | /21/ | *yau* | 'oil' | [lower][lo][-contour] |
| f. | low-rising | /23/ | *ngo* | 'I, me' | [lower][lo][+contour] |
| g. | low-level | /22/ | daai | 'big' | [lower][hi][-contour] |

In Cantonese then, Barrie suggests all tones that are [+contour] will be rising if they are specified with a [lo] value for the pitch feature, and falling if they are specified with a [hi] value for pitch.

It should be noted that although the system functions reasonably well in accounting for inventories of the languages investigated, Barrie's system has little to say about the existence of complex contour tones (e.g. /214/ in Beijing Mandarin), nor does his system, as currently implemented, seem to have a means of specifying them. In addition, although he argues for emergent, rather than universal features, the specific features postulated in Barrie's analysis are motivated by a set of implicational universals on tonal systems. The desire to ground features in universals has clear motivations, but has tended to run into difficulties with tonal systems (see §2.2.4), and I eventually argue against Barrie's proposed feature constellation for Hong Kong and Guangzhou Cantonese (in §4.2). Barrie's system is of chief interest to the present study because of its grounding in the theory associated with the contrastive hierarchy (see §2.3.3), which opens up numerous avenues for the analysis of tonal systems and allows for the possibility of a unified theoretical model which relates a given set of features to phonological activity as well as inventories, in tonal systems as well as in segmental phonology.

### 2.3.2 Cluster tones

The advent of autosegmental phonology (Goldsmith 1976) presented a substantial innovation towards the analysis of tonal systems. Based on the analysis of a number of African tonal systems, Goldsmith (1976), and Leben (1978), motivated a cluster model of tone, whereby contoured tones are concatenated composites of two or more level tones. In a language with a high rising and a low falling tone, for example, the high-rising tone could viewed as comprised of a 'medium' flat tone and a 'high' flat tone, whereas the low falling tone would be comprised of a 'medium' flat tone and a 'low' flat tone, as below:

(4) Representation of two contoured cluster tones (adapted from Barrie 2007)


As discussed in Barrie (2007), this model runs into two main areas of difficulty when applied to Chinese languages. Firstly, it tends to drastically overgenerate the number of possible contour tones, 6 for a language with a three-way height distinction, and 12 for a language with a four-way height distinction (e.g. Cantonese). In addition, if complex contour tones in African languages with a three-way height distinction do, in fact, result from the concatenation of three flat tones, we also end up with the system generating significantly more contoured tones than have been observed by surveys of tone languages (e.g. Zhang 2002), and the system does not actually ever predict the correct number of contoured tones in an inventory from the number of height distinctions the language in question makes. Secondly, and much more significantly, is the general problem discussed in depth by Yip (1989), who illustrates that the requisite individual flat tone primitives which are supposed to exist do not actually act independently in the phonology. In African languages there tends to be a one-to-one, left-to-right association of level tones, with excess tones, if associated at all, associated with the final tone-bearing unit (e.g. Leben 1973;

Goldsmith 1976; Pulleyblank 1986). Flat tone primitives are thus able to be redistributed, for example in Mende, which possesses two robust tone mapping rules:

(4) Mende tone mapping (Leben 1973)

i. If the number of level tones in a domain is equal to or less than the number of (syllables), associate the first tone with the first syllable, the second with the second, and so on; any remaining syllables receive a copy of the last tone in the pattern. Thus $^{LHL}[\textit{nikili}] \rightarrow {}^{L}[\textit{ni}]_\sigma\ {}^{H}[\textit{ki}]_\sigma\ {}^{L}[\textit{li}]_\sigma$

ii. If the number of level tones in a domain is greater than the number of (syllables), associate the first tone with the first syllable, the second with the second, and so on; remaining tones are expressed as a contoured sequence on the final syllable. Thus, $^{LHL}[\textit{nyaha}] \rightarrow {}^{L}[\textit{nya}]_\sigma\ {}^{HL}[\textit{ha}]_\sigma$

The result is that, in a given polysyllabic domain, contour tones, when present in the output, only appear on the final tone-bearing unit, regardless of the underlying representation of each unit (however see Dwyer 1978 and Conteh et al 1983 for alternate views of this phenomenon).

Yip (1989) points out that in Chinese languages, contour tones do not separate into the supposed flat tone primitives which would be required under this type of analysis, full contours generally remain associated with a given syllable. Responding to critiques such as Yip's (1989), Duanmu (1994) attempts to motivate cluster tones, arguing that in Mandarin, the tone bearing unit is the mora, thus word-internal contours are expected, and presents an analysis in which stress assignment and tone association domains are the same, pointing to the existence of unstressed, toneless syllables in Mandarin. However, these are a closed class of functional morphemes, and the same analysis would not work in, for example, Cantonese, which also permits word-internal contours, unless we postulate that every single syllable is stressed.[9]

In a further argument against cluster tones, Barrie (2007) reflects on evidence presented by Wan and Jaeger (1998) which suggests that speech error data in Mandarin can be explained by a unitary but not a cluster analysis of tone.[10] Wan and Jaeger find no tone splitting or hybridization errors resulting from blending or word telescoping. There were no cases in the study of substitution errors from replacing a supposed primitive for another primitive in any part of a word, and all errors were the result of substituting an entire tone for another.

### 2.3.3 The contrastive hierarchy

The present study builds upon the framework associated with the Contrastive Hierarchy (Dresher 2009). A central principle of this framework is the Contrastivist Hypothesis (Hall 2007) which states that only contrastive features are phonologically active. The notion of contrast derives from Saussure (1916): "dans la langue il n'y a que les differences sans termes positifs" ("in

---

9 Nantong, a Mandarin language discussed in §3.2.6, may superficially appear to be a case where an analysis such as in Duanmu (1994) works in cases of tone spreading, however, this is likely a phonetic effect in connected speech (see Ao 1993). Wu languages as a class commonly exhibit this tendency. Whether Wu languages are best analyzed this way is beyond the scope of the present study.

10 Another non-feature based model of contour tones as contoured segments is proposed by Sagey (1986). Extensive arguments against such an analysis can be found in Duanmu (1994) and Barrie (2007).

language there are only differences without positive terms"), and became central to pioneering work in Prague school phonology, in turn proving influential to many schools of thought in generative phonology (see Dresher 2009:76-136 for a detailed discussion). Dresher (2009) identifies several diagnostics which have been utilized in this tradition in the identification of contrastive features, and which, taken together, comprise an instructive, though not exhaustive, set of examples of what constitutes a contrastive feature:

(5) Diagnostics used in identifying contrastive features (Dresher 2009:72)

A phoneme φ has contrastive feature F if:

a. φ enters into an alternation or neutralization that is best explained if F is part of φ.
b. φ causes other phonemes to alternate or neutralize in a way that is best explained if F is part of φ.
c. φ participates in a series with other phonemes, Φ, with respect to phonotactic distribution, where F is required to characterize Φ in a general way.
d. the set of allophones which make up φ all have F in common.
e. speakers adapt a sound from another language in a way that can be explained by supposing that they assign F to the foreign sound.

Contrastive features are conceptualized as being arranged hierarchically, and are assigned in a given language via the *Successive Division Algorithm.* The algorithm begins with an initial state, where all sounds comprise a single phoneme, which is then subdivided into two subsets when we find evidence for the activity of a given distinctive feature. This process is repeated, successively applying features in turn until each set has only one member.

Consider, for example, a typical five-vowel system, /i e a o u/. In the initial state, all five vowels are members of a single set. If we determine the activity of a feature, say [high], we subdivide this set into two groups: [+high], which comprises /i u/, and [-high], which comprises /e a o/. We then, upon determining the activity of [back], subdivide our first group such that each has a single member, and subdivide our second group accordingly. Finally, upon determining, perhaps, that [round] is also active, we further subdivide our remaining group such that each group now has a single member, as in (6):

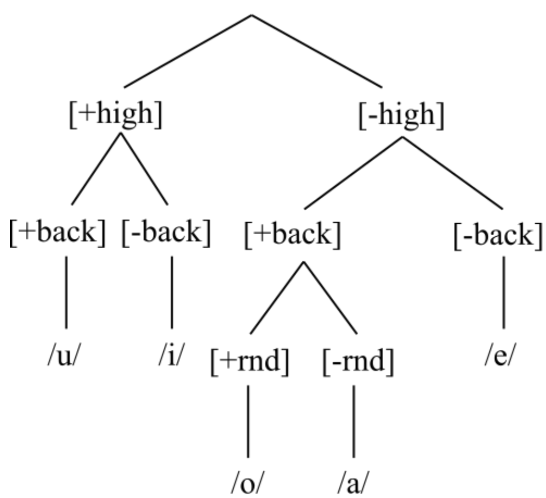
(6) [high] > [back] > [round]

/i/ [+high, -back] /u/ [+high, +back] /e/ [-high, -back]

/a/ [-high, +back, -round] /o/ [-high, +back, +round]

This can be represented hierarchically, as below:

(7) [high] > [back] > [round]




This is but one possible ordering of features, giving rise to a specific pattern of connections, oppositions, and underspecification. In actual practice, this pattern is language specific, and empirically motivated, depending upon activity. If, for example, our language exhibits a rounding process involving [+round], and only /o/ acts as a trigger, we have evidence for a specification such as in (6), whereby [high] > [back] > [round]. If, however, we note that /u/ is also acting as a trigger, we have evidence that /u/ is specified for [+round], and we would assume then that [round] has greater scope in our hierarchy, as in (8) below:

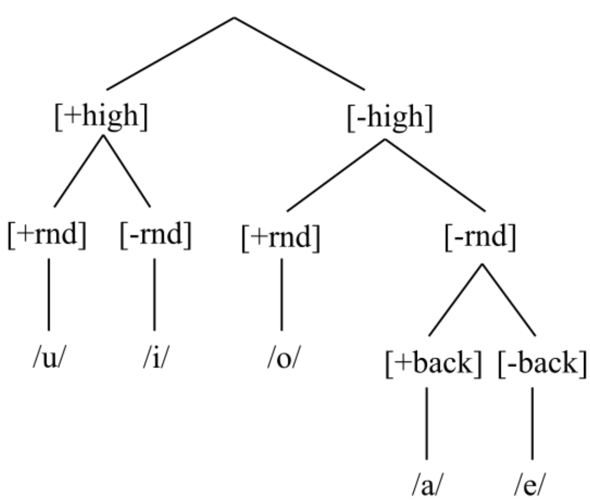
(8) [high] > [round] > [back]

/i/ [+high, -round] /u/ [+high, +round] /e/ [-high, -round, -back]

/a/ [-high, -round, +back] /o/ [-high, +round]

This is thus represented hierarchically as below:

(9) [high] > [round] > [back]




We assume that the ranking of contrasts then, in addition to the specific contrasts invoked by features, is a source of crosslinguistic variation.  As alluded to above, just as hierarchies are conceptualized as being language-specific, the content of features themselves is also emergent, rather than being universal and innately specified. In (6-8) above therefore, although we refer to features such as [high] and [back], these labels are in many ways purely notational, referring descriptively to a class of sounds and an associated class of phenomena without necessarily making criterial predictions regarding their precise manifestation in a particular language. The view is that features are defined both negatively, as a “point in a pattern” (Sapir 1925) of similarities and oppositions, and positively, possessing some amount of substance. Phonetic naturalness results when an inventory or subset thereof is in phonetic alignment with its overall pattern of contrasts, yet inventory members may also deviate from it. Features are assumed here to be binary. That is, both the positive and negative values of the feature in question are assumed to be computationally present and available to the phonology.[11]

The contrastive hierarchy resembles in some respects the framework associated with feature geometry (Clements 1985; Sagey 1986; Clements and Hume 1995 amongst others). While “earlier theoreticians tended to think of phonemes as unstructured sets of features” (Clements and Hume 1995) feature geometry, like the contrastive hierarchy, encodes structural relations between features.[12] Feature geometry is an extension of autosegmental phonology (Leben 1973; Goldsmith 1976), which noted the relative autonomy of tone with respect to segmental features. This observation provided the foundation of a theory where features are organized hierarchically, with

---

11 While privative features are in principle more restrictive, they are not necessarily more parsimonious than binary ones. This is taken up in more detail in §5, and specifically in footnote 87.

12 Dresher (2009:130) points out that Jakobson and early generative phonology experimented with a hierarchical approach to features.

certain features having scope over others, and we can think of the relative autonomy of a given subsystem as derivative of its place in a hierarchically organized constellation of features. Feature geometry as originally conceived exhibits a number of differences with the contrastive hierarchy: the presence of class nodes in addition to feature nodes, the fact that feature geometry exhibits *n*-ary rather than binary branching nodes, and the fixed and universal nature of feature geometry versus the language-specific structuring of contrasts as determined by activity in the contrastive hierarchy.[13] The last of these is worth underscoring here because assuming that features are drawn from a universal set and feature ordering is fixed leads a number of researchers to assume that tonal systems are inherently unsuited to a feature-based analysis (e.g. Clements et al. 2010; Hyman 2010, 2011).

### 2.3.4 Some arguments against tonal features

Some recent arguments against an analysis of tones as unitary entities comprised of features also exist. Such critiques tend to rest on a number of interrelated issues, but have in common at their core an inherent tension between the desire to motivate a small closed set of universal features, coupled with the fairly high degree of diversity in the phonological activity of tonal

---

[13] Dresher (2009:136) assumes that it is possible that feature geometry may be subsumed into the contrastive hierarchy.

systems and the relative autonomy of many tonal processes (Clements et al. 2010; Hyman 2010, 2011).

Clements et al. (2010) argue that, since Jakobson et al. (1952), it is usually held that a single set of features largely satisfies four functions in segmental phonology:

| Function | | example (segments) |
|---|---|---|
| Distinctive | distinguish phonemes/tonemes | /p/ and /b/ are distinguished by [±voice] |
| Componential | define correlations (sets distinguished by one feature) | [+voice] /b d ɟ g/<br>[-voice] /p t c k/ |
| Classificatory | define natural classes (rule targets, contexts) | [-sonorant] sounds are devoiced word-finally |
| Dynamic | define natural changes (e.g. assimilation) | obstruents become [+voiced] before [+voiced] consonants |

Table 3: Functions of segmental features (adapted from Clements et al. 2010)

The crux of the argument is that while it is clear that a small closed set of features performs these tasks in segmental phonology, it is unclear whether the same can be said of tonal features. Tonal inventories and alternations tend to be comparatively diverse, making a small universal set of features which accounts both for inventory membership and phonological processes difficult to postulate without adding complexity. In addition, tonal systems exhibit processes which often are unaffected by and do not interact with segmental features. Therefore, the argument goes, tonal systems must be subsumed by a distinct system.  Clements et al. state, for example that, with very few exceptions, linguists “have maintained that features are universal… thus the feature [labial] is used distinctively to distinguish sounds like /p/ and /t/ in nearly all languages of the world” (2010).

This, however, makes two critical assumptions: that the specific content of a given feature itself is universal, and that languages make contrasts in universal, rather than language-specific ways. While linguists commonly utilize terms such as [labial] or [high], utilizing these terms analytically is (or should be) distinct from postulating that these features are precisely identical in each language. Just as high front unrounded vowels are not phonetically identical in every language, the exact value of [high] is not necessarily identical in every language, either. While Clements et al. do not postulate innate features wired in via Universal Grammar, instead arguing that they arise from "common characteristics of human physiology and audition", in arguing that the diversity of tonal inventories makes a small set of universal features problematic, they seem to tacitly assume that, just because it is notationally convenient to represent similar vowel inventories with a common feature such as [high], that this feature is the 'same' in each language. Similarly, while [coronal] features commonly trigger palatalization, for example, not all [coronal] features do: just as features of a given language do not determine the precise phonetic content of inventory members in exactly identical ways as other languages, features, while often exhibiting commonalities, do not universally trigger the same processes in all languages. While it is notationally convenient to use a single label (e.g. [high] or [coronal]) it is important to realize that these refer to *classes* of analogous phenomena. That tonal systems exhibit greater variability is likely a quantitative rather than qualitative difference.

Clements et al. (2010) go on to argue that tones do not lend themselves to a feature-based analysis, perhaps due to the monodimensional nature of tone: while segments are "defined along many intersecting phonetic parameters (voicing, nasality, etc.)…tones are defined along a single parameter, F0; there is no acoustic evidence for intersecting phonetic dimensions in F0-based tone systems" (Clements et al. 2010:20). Therefore, they argue, it is difficult to see how a *universal* tone feature analysis could emerge unless wired-in by Universal Grammar, and that the best arguments for tonal features have been, on close examination, phonetically arbitrary, idiosyncratic to one language, or noncomprehensive. This observation however reinforces our previous point that the difference is likely in degrees rather than absolutes. The fact that tones are monodimensional, being defined along a single parameter, simply necessitates more variation within that parameter to make the same number of contrasts as for example, a phoneme which is defined with reference to multiple parameters. Again, problems only arise when language-specific means of encoding features are denied.

As a case study, Clements et al. propose a rough synchronic typology of East Asian tonal languages, recognizing four major types: type I, which exhibit syllable-initial voicing contrasts; type II, which exhibit distinctive phonation registers; type III, which exhibit distinct tone registers; and type IV, which possess none of the above characteristics.[14] As register features are the least controversial features proposed for tonal systems, and type III languages, they argue, are the only tonal languages which exhibit pure tonal contrasts, they test a prototypically type III language, Cantonese, to attempt to gather evidence for the phonological activity of tonal features. They propose the following inventory:

---

[14] Although they concede that this typology is over-simplistic, and further types and subtypes likely exist, they propose that most Mandarin languages fit into the fourth type. However we argue below that most Mandarin languages do seem to possess height features, and also syllable-initial voicing contrasts, despite Clements et al. (2010:14-17) predicting that such a configuration should be absent.

(10) Middle Chinese Category

| Register | I | II | III | IV |
|---|---|---|---|---|
| High register | [53] ~ [55] | [35] | [44] | [5q] [4q] [15] |
| Low register | [21] ~ [22] | [24] | [33] | [3q] |

They state that "the crucial question for our purposes is whether or not Cantonese 'activates' register distinctions in its phonology…that is, is there evidence for a feature such as [±high register] in Cantonese in the form of rules, alternations, etc.?" (2010:17). They go on to argue that Cantonese possesses a rich system of tonal alternations, citing two main types, 'changed tones' (known elsewhere in the literature as *pinjam*), and a regular tone sandhi rule. They argue that any feature-based analysis for changed tones would require a complex analysis with otherwise unmotivated "housekeeping rules", but, disappointingly, do not go into more detail. However, *pinjam* are generally treated, not as phonological processes at all, but as morpho-lexical or discourse based (see Bauer and Benedict 1997; Jurafsky 1988; Yu 2007 for a detailed discussion), and we do not expect features to play a rule in triggering this process. Clements et al. go on to investigate the regular sandhi rule in Cantonese, where the falling tone is realized as flat before a second tone of the same height, but argue that any analysis would not actually provide evidence for phonological activity of a register feature because [53] and [55] are "variants of the same tone" (2010:18). However, this is likely a simplification.[16] In modern Hong Kong Cantonese, there is

---

15 Throughout this work, a lower case –q will be suffixed to tone letters to indicate a 'checked' or short tone. Many contemporary Chinese languages make a distinction between full tones, falling on sonorant-final syllables, and checked tones, falling on obstruent-final syllables.

16 We discuss Cantonese in depth in §4.2 below.

only a single high tone, /55/, no falling tones, and no sandhi rule. In most speakers of Guangzhou Cantonese, however, there are two high tones, [55] and [53], and the sandhi rule as described above links them, as evidenced by the following minimal pair vs. homophones:

| | | |
|---|---|---|
| (11) | *sa:n*, /55/ 'hill' | *sa:n*, /53/ 'to close' (Guangzhou) |
| | *sa:n*, /55/ 'hill' | *sa:n*, /55/ 'to close' (Hong Kong) |

They then state that, in order to make a clear case for phonological alternation, a set of alternations between contrastive tones would be required, such as between /55/ and /35/. It is disappointing that Cantonese was chosen as a case study, with phonological processes set as the criteria to demonstrate the phonological activity of a height feature. As I discuss throughout §4, Yue languages as a class exhibit very little in the way of tonal processes at all,[17] and the Mandarin group, for example, would have presented numerous opportunities for an interesting analysis of phonological activity.

With specific regard to the diachronic development of Chinese tonal systems Clements et al. state that "the long history of phonetic evolution…has tended to destroy the original phonetic basis of the tone classes… this has frequently led to synchronically unintelligible tone systems" (Clements et al. 2010:19), and, citing Chen (2000:81-82), argue that the vast assortment of tonal alternations defy classification and explanation, that in one dialect after another one is "left with the baffling impression of random and arbitrary substitution of one tone with another without any

[17] In fact, Guangzhou Cantonese is the only widely reported Yue language with any robust sandhi, and it exhibits only the single process.

apparent articulatory, perceptual, or functional motivation". Again, however, this idiosyncratic behavior of tonal systems is only problematic if we assume a small closed set of universal features, or a fixed, universal means of determining contrasts. Although the diachronic changes in many languages descended from Middle Chinese are indeed complex, and there is often significant tonal drift, this should not be a problem for a synchronic analysis. In fact, as I demonstrate below, assuming lexical tones are encoded with features presents significant opportunity for insight, including regarding the diachronic development of tonal systems. For example, I show that the contrastive hierarchy of Middle Chinese is still either fully present, or exhibits traces of its former structure, in nearly all modern Mandarin (§3.2.3 - §3.2.7) and Yue (§4.2 and §4.3) languages despite the existence of significant tonal drift, and that restructuring of this hierarchy is generally predictable via phonetic (§3.2.4, §3.2.6) and general complexity (§5.4) criteria.

In a similar vein, regarding the relative diversity of tonal systems, Hyman (2010: 71) discusses the realization of an underlying /L-HL-H/ sequence in several Grassfields Bantu languages:

| (12) | <u>Language</u> | <u>Output</u> | <u>Process</u> |
|---|---|---|---|
| | Mankon | L-H-↑H | H-upstep |
| | Babanki | L-M-H | HL-fusion |
| | Babadjou | L-H-↓H | H-downstep |
| | Dschang | L-↓H-H | HL-fusion, downstep |
| | Kom | L-M-M | H-lowering |
| | Aghem | L-H-H | L-deletion |

He argues that although each language simplifies the HL portion of the input, and all but Aghem preserve a trace of both the H and the L, each language makes different choices as to what to preserve in terms of the syntagmatic relations. While interesting, this is only a problem for a feature-based analysis of tone if we assume some sort of universal constraint on activity. Instead, one might account for these differences either through a different set of contrastive specifications for each language, or simply a disparate set of relevant processes. Indeed any argument against a feature-based analysis based solely on the diversity of tone must eventually fall flat in this regard. Apparent problems only surface if we deny the possibility of language-specific means of making contrastive distinctions, for example by postulating a small, universal, closed set of features which performs in an identical manner in each language (emphasis added):

> Tone is different because of its greater diversity and autonomy compared to segmental phonology. Because of its diversity tone is hard to reduce to a ***single set of features that will do all tricks***. (Hyman 2010:69)

This is a similar argument to the one by Clements et al. (2010) against tonal features on the basis of diversity: since a single set of segmental features satisfies all the functions of a segmental phonology, and since the basis of such for tonal systems is not altogether clear, they must be subsumed by some other system that is distinct from that used to make distinctions in segmental phonology. This, however, is not necessarily a strong critique unless we postulate innately specified features which are hard-coded into the phonology. If we do suppose innate features, the critique seems to stand: a universal feature set is coded into the phonology; tonal systems do not utilize this set of features; therefore tonal phonology is not subsumed by this system. If features, however, are emergent and language specific, observed universals are not explained through some innately-specified feature set, but through factors such as “common characteristics of human physiology and audition” (Clements et al. 2010), and we can account for tonal diversity.

Arguments against features for tone based on the relative autonomy of tonal systems with respect to the rest of segmental phonology, to me, fail for some of the same reasons spelled out above. Hyman (2010:69) writes that:

> Because of its autonomy, feature systems that have been proposed, even those which relate tones to laryngeal gestures, are not reliable except perhaps at the phonetic level. Given that tone is so diverse and so poorly 'gridded in' with the rest of phonology, it is not a good candidate for universality.

To start with, the argument based on autonomy does not, in and of itself, constitute direct evidence against a feature-based system. A relative paucity of interactions between tonal features and segmental features does not imply a different computational substrate. There is a relative paucity of interactions between, for example, [round] and [coronal] features, but this is not due to each residing within their own independent systems, but rather due to the configuration of the vocal tract and the intimate connection between consonantal place of articulation and neighbouring vowel quality. Hyman (2010) argues, citing Chomsky and Halle (1968), that [+high, -low], for example, not only defines a class of vowels, but also a systematic intersection with palatal and velar consonants, whereas [+upper, -raised], for example, defines only a high tone, and not a cross-linguistically consistent class of tones. In addition, he argues, tone-laryngeal interactions are notably unreliable (2010:16), citing the case of Athabaskan languages where the same laryngeal source can correspond diachronically to either H or L tones (as in Kingston 2005).[18] However, the fact that there is a relative lack of fixed, predetermined interactions between tonal features and segmental features does not imply the absence of interactions altogether. Songjiang, in example (1) above, for example, has a direct correspondence between tone height and syllable-initial voicing, and Middle Chinese, as discussed in Section 3.3, underwent a number of tonal splits that were conditioned by initial voicing and consonancy (also see Chen 1976 for a detailed description). The common contrast across numerous Chinese languages of checked vs unchecked tones often interacts in interesting and intricate ways with syllable-final sonorancy. The argument that there is a relative lack of relationships between tonal and segmental phonology is likely true but seems to be based on a difference of degrees rather than absolutes, and this is especially true when we permit language-specific contrasts.

## 2.4 Summary

In this chapter I laid some groundwork for discussing tone in Chinese languages. I briefly reviewed different Chinese language groups and discussed Middle Chinese tonal categories. I then

---

[18] Kingston (2005) describes a fairly complex scenario whereby there are two classes of these 'mirror-image' correspondences: those which evolved from different dialects of Proto-Athabaskan (e.g. Navajo and Chipewyan), and those in which the tonal reversal occurred later. In the latter case, this was enabled by a combination of vocalic coarticulation with glottalic, stem-final consonants which were later lost, leading to creaky or tense voice vowels which subsequently were reanalyzed as low and high, respectively, and stem tones being mistaken as prefix tones in cases of reduction. These cases were then analogically extended to all tones with the same value.

discussed different models of representing tone, and proposed that the framework associated with the contrastive hierarchy possesses a number of theoretical tools which can disentangle some of the issues traditionally associated with an analysis of tonal systems. I also reviewed some recent arguments against tonal features, identifying how some of the issues treating tone with features disappear if we permit language-specific contrasts and ground contrast in activity.

# 3

# The Mandarin-Jin family and Middle Chinese

## 3.1 Overview

This chapter investigates a number of Chinese languages. Four are from the Mandarin family (Beijing, Tianjin, Nantong, and Yantai) and one is a Jin language (Pingyao). Middle Chinese split sometime after the sixth century into multiple major families, including the Wu languages (including modern Xianzhou and Taizhou), the Yue languages (which include Cantonese, Guangzhou, and Taishanese), and the Guan languages. The latter of these, Guan, forms the major family from which both of the Mandarin and Jin subfamilies of languages stem. Beijing Mandarin forms the phonological basis for modern Standard Mandarin, which is spoken more or less universally in China as a language of education and official communication. The regional distribution of each of the Guan languages studied here is given below:

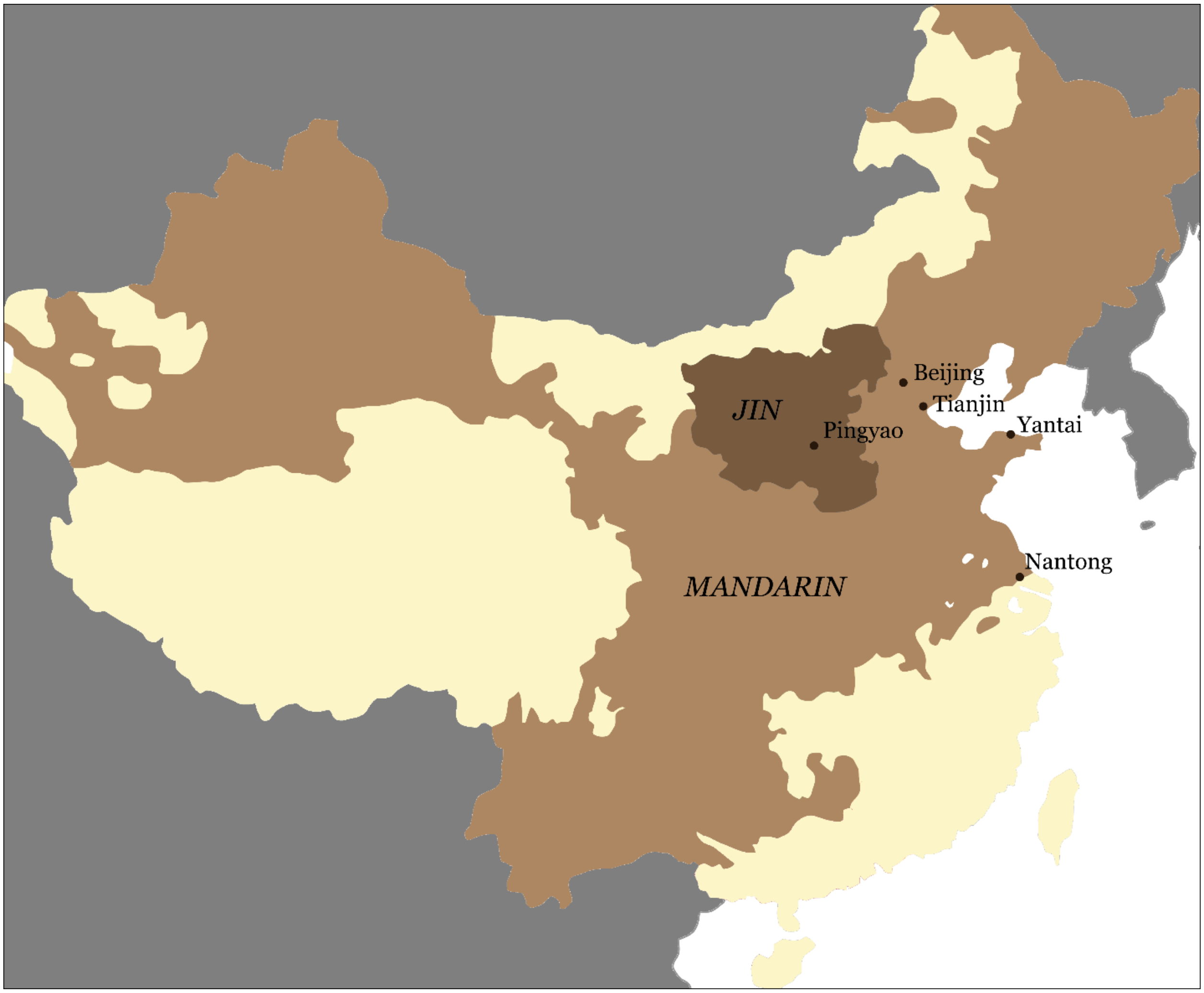



Figure 1: Geographical distribution of Mandarin and Jin languages.

## 3.2 Activity and phonetic content in tonal systems

In the following sections, we will attempt to construct a feature hierarchy for each of the Mandarin-Jin languages investigated here. As discussed in §2.3.3, this thesis assumes the Contrastivist Hypothesis (Hall 2007), "the phonological component of a language L, operates only on those features which are necessary to distinguish the phonemes of L from one another". If it is indeed the case that "only contrastive features are active in the phonology" (Dresher 2009:74), then it follows that all phonologically active features are contrastive. The Mandarin-Jin languages, as a class, exhibit an unusually high degree of phonological activity, which provides a useful starting point of analysis: the present section takes the approach of initially investigating only activity to motivate the construction of hierarchies, after which phonetic feature labelling will be applied to categories. The rationale is that if tonal systems are subsumed by the same computational system as subsumes segmental phonology, then, at a minimum, phonological activity will be subject to the same constraints and exhibit analogous patterning. As discussed above, due to the remarkable diversity inherent in tonal systems, researchers have had difficulty in devising a universally applicable set of tonal features. The present approach will serve to ameliorate this difficulty.

The following sections will provide a description and analysis of a number of Mandarin-Jin languages. Section 3.2.1 focuses on Beijing Mandarin and §3.2.2 on Pingyao. Section 3.2.3 discusses the contrastive specifications and categories of Middle Chinese as it pertains to the present discussion. Section 3.2.4 deals with Tianjin, §3.2.5 with Yantai, and §3.2.6 with Nantong. Section 3.2.7 discusses diachronic processes from Middle Chinese in more depth. Section 3.3 is a brief summary.

### 3.2.1 **Beijing Mandarin**

Let us first consider the tonal inventory of Beijing Mandarin (adapted from Chen 2000:30):

(13) Beijing Mandarin Tone inventory:

| | | | |
|---|---|---|---|
| i. | *ma* | /55/ | 'mother' |
| ii. | *ma* | /35/ | 'hemp' |
| iii. | *ma* | /214/ | 'horse' |
| iv. | *ma* | /51/ | 'scold' |

Beijing Mandarin possesses a high flat, a rising, a low concave, and a high falling tone. In certain contexts, such as reduplication, a fifth, 'light' or 'checked' tone also emerges as below:

(14) Reduplication and light tones in Beijing Mandarin:

| | | | | | |
|---|---|---|---|---|---|
| i. | *ma* | *ma* | | *ma* | *ma* |
| | 'mother' | 'mother' | → | 'mother' (address term) | |
| | */55/* | */55/* | | *55* | ***3*** |
| ii. | *bai* | *bai* | | *bai* | *bai* |
| | 'white' | 'white' | → | 'wasted effort' | |
| | */35/* | */35/* | | *35* | ***3*** |
| *iii.* | *jie* | *jie* | | *jie* | *jie* |
| | 'sister' | 'sister' | → | 'sister' (address term) | |
| | /214/ | /214/ | | 214 | **4** |
| *iv.* | *xie* | *xie* | | *xie* | *xie* |
| | 'thank' | 'thank' | → | 'thank you' | |
| | */51/* | */51/* | | *51* | ***1*** |

When any of our four citation tones in Beijing Mandarin is reduplicated, the second instance of the tone emerges as a checked, shorter, flat tone whose phonetic pitch level is predicted by the underlying tone. There is evidence that these tones are allotonic variants of their full tone counterparts, rather than actual categories in their own right: the distribution of light tones is restricted to a closed set of specific word-formation processes, and in no case does a light tone occur on an otherwise independent lexical item, thus light tones never serve as inputs into phonological processes.

Beijing Mandarin possesses two robust sandhi processes,[19] as follows:

(15) Beijing Mandarin Sandhi rule 1 (‘third tone sandhi’)

/214/ ⟶ 35 / __ /214/

| | | | | | |
|---|---|---|---|---|---|
| i. | *xiao* | *gou* | | *xiao* | *gou* |
| | 'small' | 'dog' | ⟶ | 'puppy' | |
| | /214/ | /214/ | | **35** | 214 |
| ii. | *mai* | *ma* | | *mai* | *ma* |
| | 'buy' | 'horse' | ⟶ | 'buy a horse' | |
| | /214/ | /214/ | | **35** | 214 |

(16) Beijing Mandarin Sandhi rule 2

/35/ → 55 / {/35/,/55/} __ T, T = any tone

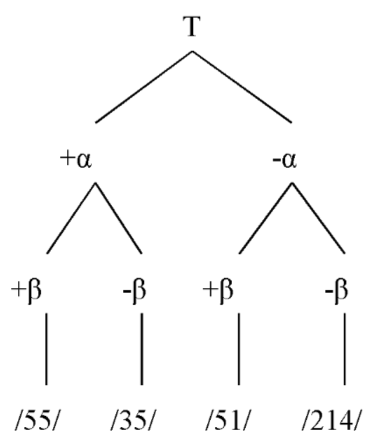
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| i. | *tian* | *wen* | *tai* | | *tian* | *wen* | *tai* |
| | 'sky' | 'hear' | 'station' | → | 'astronomical observatory' | | |
| | /55/ | /35/ | /35/ | | 55 | 55 | 35 |
| ii. | *ren* | *min* | *bi* | | *ren* | *min* | *bi* |
| | people | citizen | money | → | 'chinese currency, RMB' | | |
| | /35/ | /35/ | /51/ | | 35 | 55 | 51 |

---

19 See Chen (2000:21). The first rule is obligatory in any sequence of two consecutive /214/ that is not subject to the rule for reduplication in (14). The second rule is optional (however, see Xu 1994 for an alternate view). There is no evidence that the rules feed each other, i.e. /35.214.214/ → [35.35.214] → *[35.55.214].

Let us first start out with the assumption, central to our contrastive framework, that tones are describable with a set of contrastive features, and that each individual tonal category has its own hierarchically arranged featural specification which accounts for its phonological activity and for the way that it patterns or clusters together with other citation tones. We start with the assumption, based on general simplicity and naturalness considerations, that processes involve the change of single features.[20] Therefore, looking first at the robust sandhi paradigm in (15), where /214/ is realized as 35, we make the assumption that tone /214/ can be realized as /35/ through the change in value of a single feature. Likewise, considering rule (16), where /35/ is realized as 55, we can conclude also that /35/ can be realized as 55 through the change in value of a single feature, yielding the following possibility:

(17) Feature hierarchy for Beijing Mandarin (abstract features):




[20] While we do not rule out the possibility of more complex formulations in which processes require the change in value of multiple features, significant evidence would need to be required to postulate these. This is also based on the assumption that activity is a type of evidence which learners utilize to construct phonological hierarchies and is explored in more depth in §5.

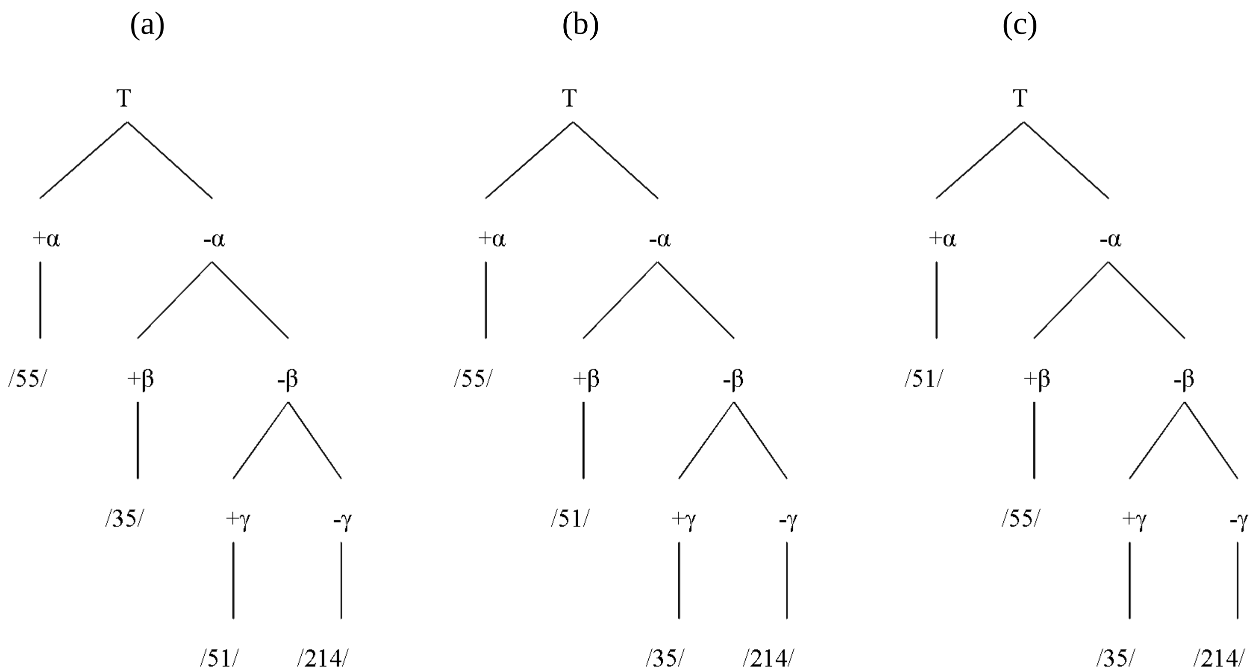
Beijing Mandarin also possesses a rule which is not a sandhi process proper, being morphologically conditioned by a small set of lexical items:

(18) Beijing Mandarin *bu-* prefixing tone change

| | | | | | |
|---|---|---|---|---|---|
| i. | *bu* | *dui* | | *bu* | *dui* |
| | 'not' | 'correct' | → | 'incorrect' | |
| | /51/ | /51/ | | 35 | 51 |
| ii. | *bu* | *dui* | | *bu* | *dui* |
| | ‘division’ | ‘squad’ | → | ‘troops’ | |
| | /51/ | /51/ | | 51 | 51 |

Example (18) illustrates the morphological nature of the *bu*-prefixing rule. When one of a closed set of four lexical items (*bu,* ‘not’; *yi*, ‘one’; *qi,* ‘seven’; and *ba,* ‘eight’), each underlyingly /51/, are attached to a root which is also underlyingly /51/, they are realized phonetically as [35]. Sequences of /51.51/ are not disallowed in any other contexts (see (18ii)), nor are there any optional rules which otherwise allow /51/ to be realized as 35. The fact that this rule is not in fact robust actually provides subtle, circumstantial evidence that a contrastive hierarchy of features constrains the activity of lexical tone. If the pattern exemplified by (18) was in fact a robust phonological rule, we would expect that /35/ can be realized as /51/ through the change in value of a single feature in this two feature system, which cannot of course be the case assuming the structure above. It is not however a robust phonological rule, but morphologically conditioned. The fact that it occurs at all suggests the possibility that a historical rule that was once robust was weeded out by a systemic change in the arrangement of contrastive specifications since it would have been otherwise incompatible with the rest of the system, leaving only this fossilized case.[21] Other technically possible but less parsimonious asymmetric 3-feature tree structures are given below:

(19) Alternative feature hierarchies for Beijing Mandarin (abstract features)[22]




Structures such as these are unlikely, however, since there isn't actually phonological evidence to prefer this type of configuration over a two-feature tree, and significant synchronic (discussed throughout §5) and diachronic (discussed in §3.2.3 and §3.2.7) evidence against it. That said, rule (15) could be accounted for in a three-feature system by assuming that /214/ becomes [+β] in (19a), which does not possess a [γ] value, or [+γ] in (19b) and (19c), and rule (16) is

---

21 At this juncture there is no evidence to decide between which of [α] or [β] constitute the first division of the tree. Regardless, there is no two-feature tree structure which will permit both of (15) and (16) as well as a robust interpretation of the *bu*- prefixing rule in (18).

22 Ignoring the individual values of features, there are only the three structurally distinct possibilities out of 12, (i.e. 4! - 2!) that permit the simultaneous existence of rules (15) and (16), assuming binary features.

likewise accounted for by assuming that /35/ becomes [+α] in (19a) and (19b) or [+β] in (19c). If this tree structure was a correct representation of Beijing Mandarin, it would still predict that the simultaneous existence of rules (15) and (16) and a robust interpretation of the *bu*- rule in (18), whereby /51/ → /35/, would be heavily dispreferred by the system in every tree but (19a), as it would entail a change of more than one feature in a single process.[23]

### 3.2.2 Pingyao Mandarin

Pingyao is a Jin dialect with a somewhat more complex set of tonal processes, possessing four citation tones, and two tones which appear only as outputs to phonological processes and do not arise independently on their own:

(20) Pingyao tone inventory (Bao 1999; Chen 2000; Hou 1980, 1989)[24]

| | | | |
|---|---|---|---|
| i. | *tçiang* | /13a/ | 'lead (the metal)' |
| ii. | *tçiang* | /13b/ | 'wall' |
| iii. | *ti* | /53/ | 'top' |
| iv. | *ti* | /35/ | 'field' |
| v. | | [31] | |
| vi. | | [423][25] | |

---

[23] It is also worthwhile to point out that the proposal that alternations generally involve tones which differ minimally constrains the possible set of alternations in a language. We would not predict the existence of a language with four tones, /a, b, c, d/, which exhibits robust bidirectional alternations between, for example, /a/ and /b/, /b/ and /c/, and /c/ and /a/ (where each tone in a pair exhibits an alternation where it can be realized as the other), as there is no possible hierarchy which permits this set of alternations assuming changes in value of single features.

[24] Pingyao also exhibits checked tone counterparts to citation tones, which occur only in closed syllables. I assume these to be allotonic with their open-syllable counterparts, following Bao (1999:75) and Chen (2000:94).

[25] Chen (2000) denotes this tone with 513.

Although merged in their respective citation forms, /13a/ and /13b/ are underlyingly distinct categories, evidenced by their differing behaviour in sandhi contexts (Chen 2000). These two tones are the modern equivalents of Middle Chinese tone I, which underwent a register split (see §3.2.7). Chen (2000:51) argues that this register split was pan-dialectal, and the citation form merger of /13a/ and /13b/ is a late, recent phenomenon. Tone /13a/ corresponds to /55/ in Beijing Mandarin, appearing on words that were formerly tone I in Middle Chinese, which possessed voiceless initials, and was likely formerly contrastively high with respect to /13b/, whereas /13b/ corresponds to /35/ in Beijing and appeared on Middle Chinese tone I words with voiced initials which were subsequently devoiced, and was likely formerly contrastively low relative to /13a/. The differing sandhi contexts for /13a/ and /13b/ in Pingyao are given below:

(21) Pingyao Tone I sandhi rule 1

/13a/ → 31 / __{/13a/ /13b/}[26]
/13a/ → 35 / /13a/__

| | | | | | |
|---|---|---|---|---|---|
| i. | *tou* | *ting* | | *tou* | *ting* |
| | ‘stealthily’ | ‘listen’ | → | ‘eavesdrop’ | |
| | /13a/ | /13a/ | | **31** | **35** |
| ii. | *tao* | *hong* | | *tao* | *hong* |
| | ‘peach’ | ‘red’ | → | ‘peach red’ | |
| | /13b/ | /13b/ | | 13 | 13 |

---

[26] In cases of reduplication, sequences of /13a.13a/ may be realized variously as [31.35] or [35.31], depending on whether the reduplicated word is a noun or a verb. Sandhi in Pingyao is often construction-sensitive and has three types: subject-predicate or verb-object; reduplicated verbs; and everything else. Structure-sensitive sandhi occurs in Jin languages, and is common in the Wu group, but is otherwise unattested in other Northern Mandarin languages.

(22) Pingyao Tone I sandhi rule 2

/13b/ → 31 / __/35/
/35/ → 13 / /13a/__

| | | | | | |
|---|---|---|---|---|---|
| i. | *tung* | *tçye* | | *tung* | *tçye* |
| | 'rainbow' | 'rainbow' | → | 'rainbow' | |
| | /13a/ | /35/ | | 13 | **13** |
| ii. | *tung* | *tçye* | | *tung* | *tçye* |
| | 'copper' | | → | 'coppersmith'[27] | |
| | /13b/ | /35/ | | **31** | 35 |

As mentioned above, the two non-citation tones, [31] and [423] appear only as outputs to phonological processes. The nature of [31] is somewhat more complex and will be dealt with below, but the other, [423] is likely simply an allotone of the citation tone /53/, as it is a target **only** of this tone and no others:

(23) Pingyao Tone II allotonic sandhi

/53/ → 423 / /35/ __

| | | | | | |
|---|---|---|---|---|---|
| i. | *xa* | *y* | | *xa* | *y* |
| | 'rain' | 'rain' | → | 'rain' | |
| | /35/ | /53/ | | 35 | 423 |
| ii. | *tung* | *xuei* | | *tung* | *xuei* |
| | 'become' | 'angry' | → | 'get angry' | |
| | /35/ | /53/ | | 35 | 423 |

---

27 The data is from Chen (2000). The gloss for *tçye* is not given.

Interestingly, this particular rule applies if it finds the appropriate environment, after all other rules have been applied, as can be seen in the output of the following rule:

(24) Pingyao Tone II dissimilatory sandhi [28]

/53/ → 35 / __ /53/

| | | | | | |
|---|---|---|---|---|---|
| i. | *er* | *ruan* | | *er* | *ruan* |
| | 'ear' | 'soft' | → | 'gullible' | |
| | /53/ | /53/ | | 35 | 423 |
| ii. | *kung* | *suei* | | *kung* | *suei* |
| | 'cook' | 'water' | → | 'boil water' | |
| | /53/ | /53/ | | 35 | 423 |

To arrive at the output in (24) we need to first apply the dissimilation rule in (24) whereby /53/ is realized as 35 before /53/, yielding [35.53], and then apply the rule in (23), yielding the final output, [35.423]. It is worth mentioning here that the Pingyao sandhi rule in (24) is the Pingyao correlate of the Beijing dissimilatory sandhi rule in (15), involving the same dissimilatory process of the same Middle Chinese tonal category. Interestingly, despite differences in input forms for this rule between Beijing, (/214.214/) and Pingyao (/53.53/), the modern output forms of this historical rule exhibit remarkable phonetic similarities: [35.214] in Beijing and [35.423] in Pingyao, despite Pingyao exhibiting a different phonological target tone (TIII in Pingyao vs. TIb in Beijing).

---

28 The data is from Chen (2000:92). The glosses for individual words are not given.

The following rule applies to /13a/, and in the output we can see yet again the effects of the rule in (23):

(25) Pingyao tone Ia sandhi

/13a/ → 35 / __ /53/

| | | | | | |
|---|---|---|---|---|---|
| i. | *tsung* | *mi* | | *tsung* *mi* | |
| | 'grind' | 'rice' | → | 'grind rice' | |
| | /13a/ | /53/ | | 35 | 423 |
| ii. | *tçi* | *ma* | | *tçi* | *ma* |
| | 'ride' | 'horse' | → | 'ride a horse' | |
| | /13a/ | /53/ | | 35 | 423 |

/13a/ is thus first realized as 35 before /53/, yielding the intermediate representation 35.53, and then the rule in (23) applies, yielding [35.423].

The following three rules are fairly straightforward, involving only a single rule application:

(26) Pingyao Tone III assimilatory sandhi[29]

/35/ → 13 / __ /13/

| | | | | | |
|---|---|---|---|---|---|
| i. | *tʰua* | *pang* | | *tʰua* | *pang* |
| | | | → | 'quit class' | |
| | /35/ | /13/ | | 13 | 13 |

---

29 The data is from Bao (1999). The glosses for individual words are not given.

| | | | | | |
|---|---|---|---|---|---|
| ii. | *xa* | *kuei* | | *xa* | *kuei* |
| | | | → | ‘start cooking’ | |
| | /35/ | /13/ | | 13 | 13 |

(27) Pingyao Tone III reduplication sandhi

/35/ → 53 / /35/ __ (reduplicated verbs only)

| | | | | | |
|---|---|---|---|---|---|
| i. | *ung* | *ung* | | *ung* | *ung* |
| | ‘inquire’ | ‘inquire’ | → | ‘to ask’ | |
| | /35/ | /35/ | | 53 | 31 |

(28) Pingyao tone II reduplication sandhi

/53q/ → 13 / /53/__

| | | | | | |
|---|---|---|---|---|---|
| i. | *ciq* | *ciq* | | *ciq* | *ciq* |
| | ‘mat’ | ‘mat’ | → | ‘straw mat’ (reduplicated noun) | |
| | /53/ | /53/ | | 53 | 13 |

Rule (28) needs to be differentiated from the rule in (24). Although both are dissimilatory rules that apply to sequences of two consecutive /53/ tones, unlike (24), (28) applies only to cases of reduplication in closed syllables.

The final rule we look at involves a simple dissimilatory sandhi where /35/ is realized as the non-citation tone [31].

(29) Pingyao tone III dissimilatory sandhi

/35/ → 31 / __ /35/

| | | | | | |
|---|---|---|---|---|---|
| i. | *pa* | *cin* | | *pa* | *cin* |
| | | | → | ‘upset’ [30] | |
| | /35/ | /35/ | | 31 | 35 |
| ii. | *tsi* | *tsa* | | *tsi* | *tsa* |
| | ‘cut’ | ‘vegetable’ | → | ‘cut vegetables’ | |
| | /35/ | /35/ | | 31 | 35 |

As we are primarily interested in analyzing patterns of activity to determine contrastive specifications, let us summarize the alternations we observe in Pingyao sandhi:

| Input | Outputs | Rule |
|---|---|---|
| /13a/ | 31, 35 | (21i), (25) |
| /13b/ | 31 | (22ii) |
| /35/ | 13, 53, 31 | (26), (27), (29) |
| /53/ | 423, 35, 13 | (23), (24), (28) |

Table 4: Pingyao sandhi tonal alternations.

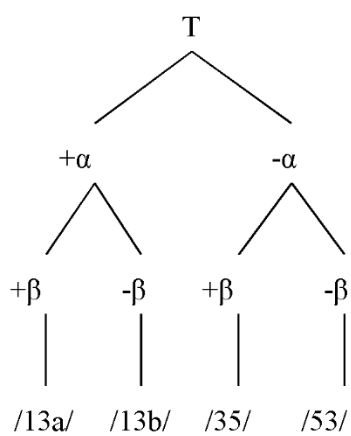
Towards determining the hierarchical organization of features for Pingyao, we can first ignore [423] as it is simply an allotone of /53/, and gives little extra information regarding activity. A possible confounding factor here is that /13a/ and /13b/ are merged in citation form, yet we have 13 as an output to processes involving both /35/ and /53/. In the absence of other information there is no a priori way to conclude which of the respective featural specifications of /13a/ or /13b/ is

30 The data is from Bao (1999). The gloss for individual words is not provided.

the target of these sandhi rules (i.e. (26) and (28)). Likewise, [31] is also not a citation form, yet is the output of processes involving /13a/ (21i), /13b/ (22ii), and /35/ (29), and we have no a priori means of determining its underlying form, so let us for the moment try to construct a feature tree without deciding these issues.

We can observe first that /35/ differs from /53/ by a single feature, and that this alternation is bidirectional (examples (24) and (27)), implying that the two differ by a single feature, and are specified for the same number of features. The observation that /13a/ can also be realized as /35/ via the change in value of a single feature (example (25)), entails a two-feature symmetric system. This is enough information to conclude therefore that /13a/ and /13b/ also differ by a single feature, which is a conclusion we might expect by default given that they originate from the same tonal category in Middle Chinese.[31] This yields the following:

(30) Pingyao (abstract features):




[31] The bidirectional alternations in (24) and (27) coupled with the rule in (25) rule out any 3-feature tree altogether. Given that there are 4 possible specifications in a 2-feature symmetrical tree, and given that /13a/ differs from /35/ by a single feature, that /53/ and /35/ differ by a single feature, and that /53/ is not equivalent to /13a/ (thus /53/ and /13a/ share **no** features), /13b/ differs from /13a/ by a single feature.

This hierarchical configuration therefore predicts that the sandhi rule in (26) should be analyzed as /35/ → /13**a**/, and the sandhi rule in (28) should be analyzed as /53/ → /13**b**/.

Let us now turn to the precise nature of [31]. Assuming that [31] is allotonic with one of our citation tones, we predict that, since the system disfavours alternations that involve a change in value of two features simultaneously, the contrastive specifications of [31] can only be the output of processes whose input differs from it by at most a single feature. Thankfully, [31] does not break our system by being the output of all four tones, but is instead the output only of /13a/, /13b/ and /35/. Out of these three tones, /35/ and /13b/ share no features, thus 31 cannot be allotonic with either of them, as this would necessitate a process involving a change of two features rather than one, and therefore 31 is analyzed as being allotonic to /13a/.

This decided, the entire range of sandhi processes involving feature changes in Pingyao are summarized below:

| Tone Change | | | Feature Change | | Rule |
|---|---|---|---|---|---|
| 13b | → | 31 (13a) | [+α, -β] → | [+α, **+β**] | (22ii) |
| 53 | → | 35 | [-α, -β] → | [-α, **+β**] | (24) |
| 13a | → | 35 | [+α, +β]→ | [**-α**, +β] | (25) |
| 35 | → | 13a | [-α, +β] → | [+**α**, +β] | (26) |
| 35 | → | 53 | [-α, +β] → | [-α, **-β**] | (27) |
| 53 | → | 13b | [-α, -β] → | [+**α**, -β] | (28) |
| 35 | → | 31 (13a) | [-α, +β] → | [**+α**, +β] | (29) |

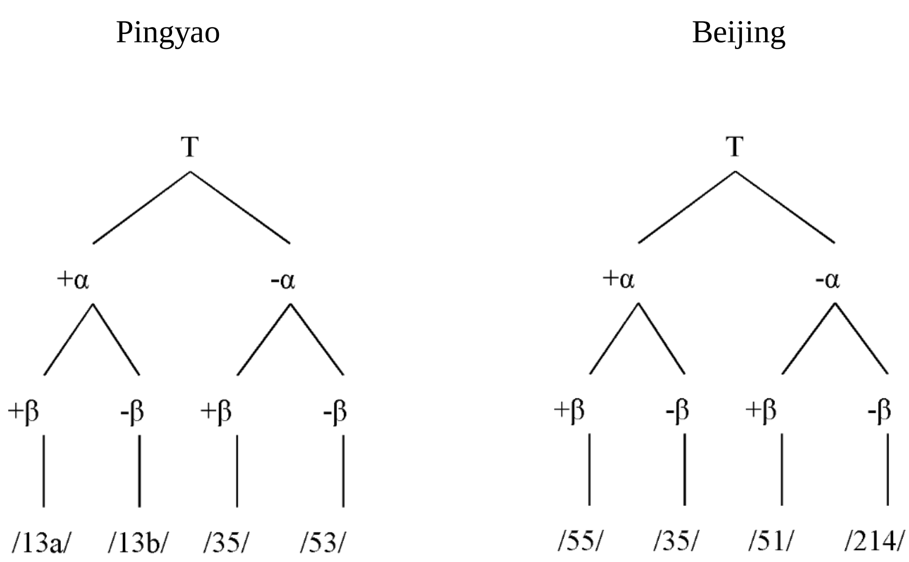
Table 5: Feature changes in Pingyao sandhi.

As with Beijing Mandarin, the Pingyao tonal inventory and sandhi rules seem amenable to an analysis based upon a contrastive hierarchy of features grounded in phonological activity. Despite the highly robust and complex range of tonal alternations in Pingyao, they all conform neatly to the prediction that rules out changes of more than single features in any alternation process.

### 3.2.3 Middle Chinese categories and abstract vs. concrete features

Before moving on to an analysis of Tianjin, let us first more closely examine the feature hierarchies of Pingyao and Beijing alongside one another, reproduced below:

(31) Pingyao and Beijing Mandarin feature hierarchies (abstract):




It was mentioned above that Beijing sandhi rule 1 (15) and Pingyao sandhi rule 2 (24) are in fact the modern correlates (i.e. a dissimilation process that occurs in sequences of two consecutive [-α, -β] tones) of the same historical rule, occurring in the same context and with the same output (that is, the first of the two tones becomes [-α, +β] ). As will be discussed in more detail in §3.3, despite the presence of somewhat drastic differences in the precise phonetic manifestation of the Beijing and Pingyao inventories, evidence from documented phonological changes and data from cognates illustrate that the tones which exhibit the same featural specifications in (31) are also the same tone categories diachronically. The modern correlates of the same Middle Chinese tonal categories possess the same structural relationships with one another in these languages and predict the possible ranges of sandhi outputs despite highly disparate sets of respective sandhi processes. This is illustrated in the following table. The Middle Chinese cognates come from Newman and Raman's (1999) reconstruction of Middle Chinese.[32] The cause for the fact that both /214/ and /51/ in Beijing, and /53/ and /35/ in Pingyao possess MC tone II as a tonal cognate is due to a partial Tone II ⟶ Tone III merger in late middle Chinese and will be discussed in more depth in §3.2.7.

---

32 For a more detailed discussion on the reconstruction methodologies see Chen (1976), Chen and Newman (1984a, 1984b, 1985), Newman and Raman (1999), and Raman (1997).

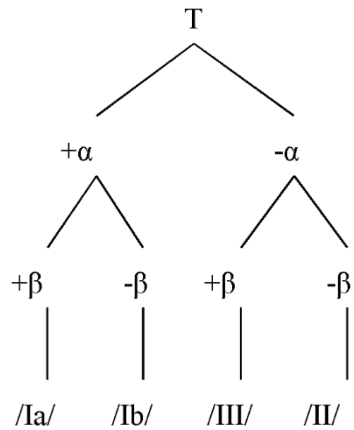
| MC Tone | MC cognate | Beijing | Pingyao | Character | English gloss |
|---|---|---|---|---|---|
| | | /55/ [+α +β] | /13a/ [+α +β] | | |
| Ia | kuɑn | guan | kuang | 关 | ‘close’ |
| | k$^h$ɑi | kai | kae | 開 | ‘open’ |
| | tso | zu | tseu | 租 | ‘rent’ |
| | | /35/ [+α -β] | /13b/ [+α -β] | | |
| Ib | duŋ | tong | tong | 铜 | ‘copper’ |
| | dziɑŋ | qiang | tçiang | 墙 | ‘wall’ |
| | diai | ti | ti | 蹄 | ‘hoof’ |
| | | /214/ [-α -β] | /53/ [-α -β] | | |
| | miai | mi | mi | 米 | ‘rice’ |
| | ma | ma | ma | 马 | ‘horse’ |
| II | | | | | |
| | | /51/ [-α +β] | /35/ [-α +β] | | |
| | ziu | shou | seu | 受 | ‘receive’ |
| | | /51/ [-α +β] | /35/ [-α +β] | | |
| III | di | di | ti | 地 | ‘field/ground’ |
| | ka | jia | tçia | 价 | ‘price’ |
| | ts$^h$ɑi | cai | tsae | 菜 | ‘vegetable’ |

Table 6: Beijing and Pingyao tonal categories and Middle Chinese cognates.

As can be seen, there is a one-to-one relationship between Beijing and Pingyao in all categories; the same featural representation in both languages corresponds to the same category in Middle Chinese. This relationship can be schematized as follows:[33]

[33] The modern tones with the specification [-α +β] represent the modern correlate of the **category** that was tone III in MC, despite the fact that the modern category also contains some members that were formerly of tone II in MC due to the partial merger of tone II with tone III described in Chen (1976, 2000), discussed in more depth in §3.2.7.

(32) Middle Chinese tonal cognates of modern Beijing and Pingyao categories




A more purely phonetic-based means of assigning features to the above patterns might involve something like the following. One could start with examining only the set of Pingyao features, where it would become immediately obvious that tones with the [+α] feature are 'low' whereas those with the [-α] are high. Recalling that [31] is allotonic with /13a/ in Pingyao, we might make the observation that, regarding our [β] feature, both [-β] tones lie at the pitch periphery of the tonal space already divided by the [α] feature: /53/ and /13b/, whereas both [+β] tones, /35/ and /13a/ and [31] lie more towards the center. One could perhaps postulate then that our [β] feature relates to a pitch being [inner] or [-peripheral]. A [peripheral] feature divides the pitch range into three registers, with positive values indicating a tone lies within either the upper- or lowermost register, and a negative value indicating a tone lies in the center.[34] Although this gives no indication

[34] Here, [peripheral] is roughly analogous to what Odden (2010) identifies as [extreme]. See Appendix B. In the absence of evidence to the contrary, we adopt Yip's (2001) proposal that height in contour tones is only specified at the left edge, and thus treat /53/, rather than /35/, as peripheral, despite both tones' citation forms possessing the same overall pitch range. In any case, the determination of the precise phonetic content of the features here remains a challenge due to the existence of two phonologically distinct yet phonetically merged classes in citation form.

of the presence of contour or directionality, it does serve as a means of contrastively dividing the four tones into their respective categories. Regarding Beijing Mandarin, although the same feature labels clearly would not apply to this language, one could certainly cobble together a language-specific phonetic description of the specifications, for example by stipulating that a [-α] value (/214/ and /51/) corresponds to a 'falling' element of the tone and a [+α] value (/55/ and /35/) corresponds to 'not falling'. One could then make the observation that both of our [+β] tones (/51/ and /55/) are high, and the other two tones (/35/ and /214/) are not. We could then conclude that in Beijing Mandarin, low tones are, by default, rising and we have a phonetic description of Beijing Mandarin.

It will eventually be shown that it is possible to arrive at some phonetically plausible, language-specific interpretation of each feature that suitably describes the tonal inventory of each of the languages in the present study. That said however, the fact that the tones at the same location in their respective contrastive hierarchies, motivated by the phonological activity of highly disparate processes, also correspond exactly to the same tonal categories in Middle Chinese suggests a deeper relationship between the internal organization of contrastive features and a tonal system than one might otherwise imagine if one took a more concrete or universalist view of features themselves. Although the structural skeleton of contrastive specifications seems remarkably durable here, diachronically speaking, the precise phonetic manifestation seems more subject to change and variability over time. In fact, the phonetic variability exhibited between Pingyao and Beijing Mandarin exemplifies some of the inherent variability in tonal systems in general that motivated researchers such as Hyman and Clements et al. discussed above, to discard a feature-based representation altogether. This deeper relationship disappears, however, if we superficially equate the abstract dimension of features as category delimiting entities within a discrete formal system, with their concrete representations.

The contrastive hierarchy thus allows insights such as the above, and this stems directly from the specifics of how features are conceptualized. In conceptualizing features as possessing both an abstract dimension consisting of a set of hierarchically arranged interrelations and oppositions, which determines the clustering of phonemes and patterning of processes, as well as a language-specific, concrete dimension containing phonetic content, we avoid many of the problems associated with pairing phonetic content or hierarchies with static universals. If we adopt the converse assumption, that features are concrete at the computational level, and every phonetic change in the tonal inventory implies a concomitant change in the phonological system, we end up with an odd consequence if we examine the respective Beijing and Pingyao hierarchies with respect to Middle Chinese. Notationally, one could change the labels attached to the features, at every point in time when a phonetic change in the tonal inventory renders the previous labels descriptively inadequate, but doing so would have absolutely no consequence in terms of the systemic behaviour: activity patterns and the relationship between categories remains the same in both languages despite individually different sandhi processes, and differences in phonetic inventories. We are left thus with the odd consequence of having supposedly phonological changes with absolutely no phonological consequence, and a theory which postulates concrete, universal features (e.g. Hale and Reiss 2008) would be unable to account for analogous activity patterning in Beijing and Pingyao stemming from a common ancestral structure despite exhibiting superficial phonetic differences.

Of course, maintaining a conceptual separation between the abstract and concrete dimensions of features, that is defining them as fundamentally dyadic, is by no means a new concept. Sapir (1925) proposed two sets of hypothetical languages, A and B, which had identical phonetic inventories but different 'patterns', or ways that the inventories are organized, and hypothetical languages C and D, which had identical 'patterns' but different phonetic inventories, reproduced below (adapted from a discussion in Dresher 2011):

(33) Phonemes with identical patterning

Pattern of C

| | | | | | | |
|---|---|---|---|---|---|---|
| h | | w | j | l | m | n |
| p | t | | k | q | | |
| b | d | | g | G | | |
| f | s | | x | χ | | |

Pattern of D

| | | | | | | |
|---|---|---|---|---|---|---|
| h | | v | ʒ | r | m | ŋ |
| $p^h$ | $t^h$ | | $k^h$ | $q^h$ | | |
| ß | ð | | ɣ | ʁ | | |
| f | ʃ | | ç | ħ | | |

Dresher argues that Sapir conceptualized a particular phoneme's point in the pattern as referring to “the contrastive status of a phoneme, and the way it relates to other phonemes in the system” (2011: 245), and despite the fact that Sapir postulated no theory of features explicitly, such a notion seems obligatory to render this conceptualization meaningful. Interestingly, Sapir (1925) goes on to argue that although a particular 'pattern' is guided by phonetics, it may also deviate from it, as in /ʒ/ in language D, which is physically an obstruent, but psychologically and functionally a sonorant, corresponding to /j/ in C and being classified with that group of sounds. It seems that in Sapir's example, a 'point in the pattern' refers to a set of abstract structural relations, which may, but do not obligatorily, exhibit an isomorphism with a set of phonetic criteria. Following Sapir's conceptualization, Trubetzkoy (1939:13) also notes that: “the linguistic values of sounds to be examined by phonology are abstract in nature. They are above all relations, oppositions etc., quite intangible things, which can be neither perceived nor studied with the aid of the sense of hearing or touch”. This view is echoed in more modern formulations regarding the abstract dimension of features, such as Blaho (2008: 22-23) who argues that features are “indicators of the way members of an inventory behave, but they don’t necessarily have any consistent phonetic characteristics... there is no reason a priori why phonological features have to correspond to phonetic properties.” Similarly, Mielke (2008:9) argues compellingly for the view that features are “abstract categories based on generalizations that emerge from phonological patterns”; while features are abstract and there is not necessarily a direct relation between phonetics and phonological patterns, features nevertheless reflect phonetics “*via* the… phonological patterns they are motivated by” (2004:9). This is in rather stark contrast to the purely innatist view, which sees phonological patterns as direct products of universal features, and thus phonetic content more or less directly reflect features as well. This is the perspective advanced in *The Sound Pattern of English* (Chomsky and Halle 1968:294): "the total set of features is identical with the set of phonetic properties that can in principle be controlled in speech; they represent the phonetic capabilities of man and, we would assume, are therefore the same for all languages".

### 3.2.4 Tianjin Mandarin: Contrastive reanalysis and phonetic change

Let us turn now to an investigation of another Mandarin language, Tianjin, which, as we will see, exhibits a more nuanced set of diachronic developments than Beijing or Pingyao. The tonal inventory of Tianjin is as follows:

(34) Tianjin Mandarin tone inventory[35] (Bao 1999; Li and Liu 1985; Shi 1990)

| | | | |
|---|---|---|---|
| i. | kɑu | /21/ | 'high' |
| ii. | fəŋ | /45/ | 'to sew' |
| iii. | ɕi | /213/ | 'to wash' |
| iv. | xuei | /53/ | 'meeting' |

Tianjin has been argued to exhibit four robust two-syllable sandhi processes (Bao 1999; Chen 2000: 105-106; Li and Chen 2012; Zhang and Liu 2011, 2016). One of these, (35), is present in most other Mandarin languages,[36] and corresponds to the rules in (15) in Mandarin, and (24) in Pingyao above, involving dissimilation of the same tonal category in the same environment:[37]

35 Shi (1990) classifies the tones as /11/, /55/, /24/, and /53/ respectively.

36 Mei (1977) traces this pattern back at least to the 16th century.

37 That is, in each language, sequences of two consecutive modern correlates of the Middle Chinese tone category /II/ are realized as a rising tone followed by a concave tone, despite phonetic differences in input forms.

(35) Tianjin sandhi rule 1 /213/ → 45 / __ /213/[38]

| | | | | | |
|---|---|---|---|---|---|
| i. | ɕyan | tsuŋ | | ɕyan | tsuŋ |
| | | | → | ‘to select seeds’ | |
| | /213/ | /213/ | | **45** | 213 |
| ii. | pau | kuan | | pau | kuan |
| | | | → | ‘to safe-keep’ | |
| | /213/ | /213/ | | **45** | 213 |

Interestingly, in all three languages, the output forms of this rule exhibit close phonetic similarity, despite obvious differences in input forms: /214 214/ → [35 214] in Beijing, and /53 53/ → [35 423] in Pingyao. The other two-syllable sandhi processes are unique to Tianjin. They were acquired after the bifurcation of Tianjin from the main Mandarin branch, and have been argued to involve dissimilation, Tonal Absorption (Chen 2000:106) or Contour Metathesis (Bao 1999:61). It will be argued here that all sandhi processes in Tianjin are simply dissimilatory in nature.

(36) Tianjin sandhi Rule 2 /21/ → 213 / __ /21/

| | | | | | |
|---|---|---|---|---|---|
| i. | *kau* | *ʂan* | | *kau* | *ʂan* |
| | | | → | ‘high mountain’ | |
| | /21/ | /21/ | | **213** | 21 |

---

38 The data from this section is from Bao (1999), Chen (2000); Li and Chen (2012); Li and Liu (1985); Shi (1990); Zhang and Liu (2011, 2016). When absent, glosses for individual words are not provided in the source consulted.

| | | | | | |
|---|---|---|---|---|---|
| ii. | *tʂu* | *fa* | | *tʂu* | *fa* |
| | | | → | 'to set out' | |
| | /21/ | /21/ | | **213** | 21 |

(37) Tianjin sandhi rule 3 /53/ → 45 / __ /21/

| | | | | | |
|---|---|---|---|---|---|
| i. | *ta* | *ku* | | *ta* | *ku* |
| | | | → | 'aunt' | |
| | /53/ | /21/ | | **45** | 21 |
| ii. | *faŋ* | *ɕin* | | *faŋ* | *ɕin* |
| | | | → | 'at ease' | |
| | /53 | /21/ | | **45** | 21 |

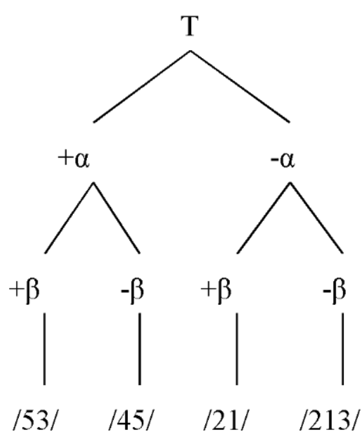
Another ostensible sandhi process has been described in the literature. Although we will argue that this no longer a robust phonological process, it is reproduced below in (38) for convenience:

(38) Tianjin sandhi rule 4 /53/ → 21 / __ /53/

| | | | | | |
|---|---|---|---|---|---|
| i. | *tɕian* | *tsu* | | *tɕian* | *tsu* |
| | | | → | 'architecture' | |
| | /53/ | /53/ | | **21** | 53 |
| ii. | *xuei* | *i* | | *xuei* | *i* |
| | | | → | 'meeting' | |
| | /53 | /53/ | | **21** | 53 |

Following a similar logic as with Beijing and Pingyao, we can deduce a feature hierarchy with the observations that /213/ and /45/ differ by a single feature, /21/ and /213/ differ by a single feature, and /53/ and /45/ differ by a single feature, yielding one possible hierarchy.[39] If (38) was an actual phonological process, it would imply that /53/ and /21/ also differ by a single feature, but this is redundant information, and the construction of our feature tree in (39) is unchanged by its exclusion.

(39) Feature hierarchy for Tianjin Mandarin (abstract features):




The clearest phonetic labelling of these would likely be [+high] for [+α] or [+low] for [-α], and [+falling] for [+β] or [+rising] for [-β].

---

39 We assume [α] > [β] here because /53/ and /45/ are the modern correlates of Middle Chinese tones /Ia/ and /Ib/ which arose from a register split of a single tone, and should thus be contrastive sisters by default without evidence to the contrary. However [β] > [α], which results in different terminal sisters, is not explicitly ruled out by any evidence thus far.

| MC Tone | MC cognate | Beijing | Tianjin | Character | English gloss |
|---|---|---|---|---|---|
| | | /55/ [+α +β] | **/21/ [-α +β]** | | |
| Ia | kɑu | gao | kɑu | 高 | ‘high' |
| | t$^{h}$ɑ | ta | t$^{h}$a | 他 | ‘he' |
| | fyaŋ | fang | fɑŋ | 方 | ‘square' |
| | | /35/ [+α -β] | /45/ [+α -β] | | |
| Ib | zi | shi | ʂi | 时 | ‘time’ |
| | jyan | rong | ɹuŋ | 荣 | ‘honored’ |
| | biəŋ | ping | p$^{h}$iŋ | 评 | ‘comment on’ |
| | | /214/ [-α -β] | /213/ [-α -β] | | |
| | siai | xi | ɕi | 洗 | ‘to wash’ |
| | t$^{h}$iai | ti | t$^{h}$i | 体 | ‘body’ |
| II | | | | | |
| | | /51/ [-α +β] | **/53/ [+α +β]** | | |
| | zi | shi | ʂi | 是 | ‘to be’ |
| | | /51/ [-α +β] | **/53/ [+α +β]** | | |
| III | ɣuɑi | hui | xuei | 会 | ‘meeting’ |
| | k$^{h}$io | qu | tɕ$^{h}$y | 去 | ‘to go’ |
| | si | si | sɨ | 四 | ‘four’ |

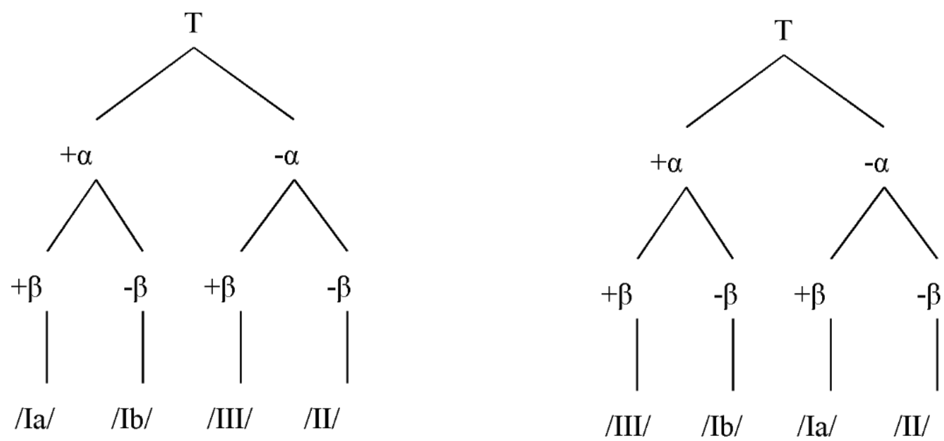
Table 7: Beijing and Tianjin tonal categories and Middle Chinese cognates.

With respect to the diachronic relation of Tianjin categories vis-à-vis those of Beijing and Pingyao, Tianjin is a somewhat more nuanced case. Although it exhibits the same tonal categories with the same historical origins as both Beijing and Pingyao, the contrastive specification of those categories in modern Tianjin is different, as can be seen in the table of cognates above.

The difference between Beijing/Pingyao and Tianjin is made clearer if we compare the feature representations of cognate categories in tree form, as below:

(40) Middle Chinese tonal cognates of modern Beijing and Pingyao vs. Tianjin

Beijing/Pingyao Tianjin



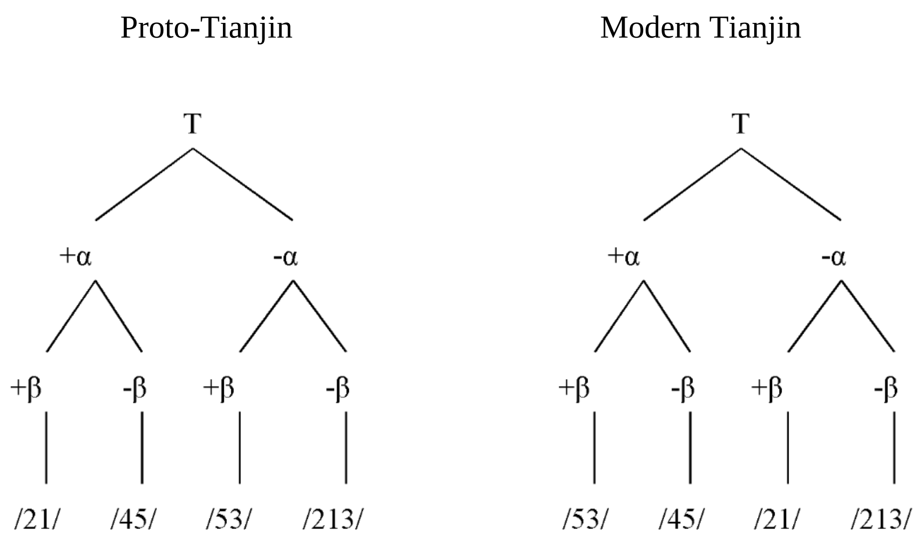

Although each tonal category in all three languages contains more or less the same words or cognates that a corresponding category in another of the three languages does, the relationships between the categories are different in Tianjin with respect to the other two. Oxford (2015) hypothesizes reanalyses of contrastive specifications such as these in a number of Algonquian languages with the *Contrast Shift Hypothesis*, and the *Segmental Reanalysis Hypothesis:*

(41) ***Contrast Shift Hypothesis****:* Contrastive hierarchies can change over time.

(42) ***Segmental Reanalysis Hypothesis****:* A segment may be reanalyzed as having a different contrastive status.

Given that there is still a one-to-one correspondence in each of the three languages between tonal categories and their cognate categories, the contrastive reanalysis in Tianjin occurred after the mergers and splits discussed above, making it reasonable to assume that the contrastive specifications of the Tianjin tonal inventory, at one point in the past, resembled those of modern Beijing and Pingyao. The question as to what conditioned the contrastive reanalysis becomes clear if we postulate a Proto-Tianjin language, exhibiting the contrastive specifications of Beijing/Pingyao with the modern Tianjin phonetic values, and compare this with modern Tianjin, as below:

(43) Contrastive reanalysis of Proto-Tianjin




In our hypothetical Proto-Tianjin, neither our two low tones, */213/ and */21/, nor our two high tones, */45/ and */53/ share features with each other, necessitating a less natural and more complex encoding of features. Tonal drift likely accreted changes in height values until the system of contrasts reached some critical inflection point which precipitated the reanalysis of specifications. If one conceptualizes systemic change as occurring with subsequent generations of learners formulating new hypotheses about a phonological system from available data, a contrast shift such as this represents a new hypothesis of a system that is more phonetically natural and possesses a lower complexity than the preceding one. Under this scenario, learners always select the algorithmically simplest possible hierarchy permitted by evidence from phonetic shape and phonological activity. This will be covered in more depth in §5.

In the vast majority of Chinese languages, including all of the Mandarin/Jin and Yue languages which are investigated in the present study,[40] the modern correlates of middle Chinese tone Ia are phonetically higher in some respect, or have a higher allotone than their Ib counterparts, except for Tianjin, where the opposite is the case. Late Middle Chinese tone Ia and Ib are the products of a register split of early Middle Chinese tone I (see Chen 1976, 2000). The register split coincided with and was conditioned by the devoicing of syllable-initial consonant voice contrasts.[41] Syllables formerly beginning with a voiced consonant (Tone Ib) were reanalyzed as having a contrastively low tone, whereas those with voiceless consonants were reanalyzed as having a contrastively high tone (Tone Ia), and the loss of a syllable initial voicing contrast coincided with the gain of a tonal contrast. This is discussed in more detail below (§3.2.7). Therefore, although we lack a detailed understanding of the precise phonetic shapes of the entire Middle Chinese tonal inventory, we can be reasonably sure that significant drift occurred in Tianjin in the modern correlates of late Middle Chinese tone Ia (modern Tianjin tone /21/), which was previously high and drifted downward, and tone Ib (modern Tianjin tone /45/) which was previously low and drifted upward.

We can thus motivate the following:

(44) ***Corollary to the Contrast Shift Hypothesis*****:** *Changes in the contrastive organization of features are motivated by phonetic change, whether local or systemic.*

---

40 With the exception of Yangjiang, a Yue language, examined in more detail in §4.3.5, whose tones exhibit an interesting and unusual diachronic trajectory.

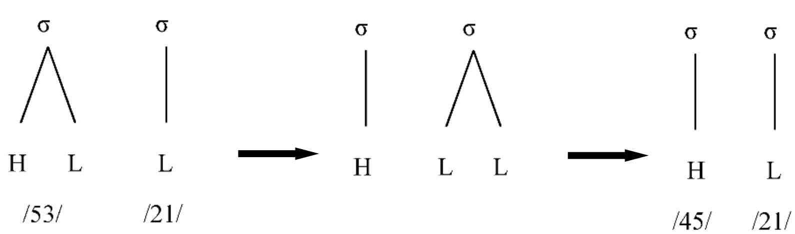
41 Research by Mei (1970) also argues that Middle Chinese tone I had two allotonic variants.

A proposal such as this is not particularly controversial. Oxford (2015), argues that "contrast shifts are influenced ... by the vaguely defined but generally-recognized forces of drift, markedness and symmetry". As is discussed in more detail in §5, I argue that such reorganization is predicted by the minimization of complexity, or algorithmic entropy, and constrained by both phonological activity and the phonetic shape of inventory members.

It is worthwhile asking why Tianjin underwent a reanalysis whereas Beijing and Pingyao did not. Certainly neither of Beijing nor Pingyao represent the most phonetically natural organization of tones imaginable. Although this will be explored in more detail in §5, the facts surrounding the diachronic and synchronic behaviour of these three do suggest some general principles. Firstly, the fact that contrastive specifications represent a fairly durable structural skeleton around which phonological activity is patterned, without direct reference to the particulars of the phonetic shape of the tonal inventory, suggests that the set of phonological processes in a particular language is at least as important a factor as the particulars of the phonetic inventory of tones when learners formulate hypotheses about a language. Beijing and Pingyao, despite possessing radically different tonal inventories, possess the same underlying structure, and although the phonetic feature labels suggested for each language above may not be prototypes for phonetic naturalness, the systems themselves still represent learnable systems. More critically perhaps, in Beijing and Pingyao, there are no clear **alternative** systems utilizing the same tonal inventories in a different hierarchical configuration that represent more natural, less computationally difficult alternatives to the present systems; it is not only the learnability of the present system that needs to be taken into consideration, but the learnability of any possible alternative systems as well. In addition, although we are examining feature hierarchies through the lens of features as abstract categories in part to underscore their diachronic origins and behaviour as categories, it should be emphasized that the features themselves still need to do the work of generating some concrete phonetic specifics of the inventory in question and that the overall hierarchical arrangement is motivated here by a set of concrete sandhi processes. Therefore, in cases such as Tianjin, in order to reanalyze a hierarchy to a less complex, more phonetically natural alternative, the alternative must already be implied by the current phonetic inventory, and the pressure to reanalyze needs to exceed the effort required to either weed out or alter the set of (older) sandhi processes that would otherwise violate the (new) alternative configuration.

Regarding motivation for two syllable sandhi processes, Chen (2000:106) argues that the rule identified in (37) is a consequence of Tonal Absorption (as in Hyman and Schuh 1974). Assuming a type of tone cluster analysis, Hyman and Schuh (1974:90) define absorption as a subtype of spreading whereby the rightmost component of a contour tone, when followed by an identical level tone, is deleted, satisfying an OCP requirement, as below:

(45) Tonal Absorption (adapted from Chen 2000; Hyman and Schuh (1974)):




In order for this analysis to work however, we need to assume that /45/ does not possess a contour (Chen 2000 classifies /53/ as F, 'falling', /45/ as H, 'high', /21/ as L, 'low' and /213/ as R, 'rising'), technically a possibility, but in the absence of evidence from activity, not a conclusion we should necessarily jump to based solely on reported phonetic shape. Indeed empirical data in Li and Y. Chen (2012) on phonetic shape of citation tones in Tianjin show a rise in /45/ of approximately 3 semitones, which is approximately the same degree of pitch drop reported for /53/, which is uncontroversially falling.[42] In addition, Wee (2010) points out that postulating this automatic process in Tianjin would actually predict the existence of two other sandhi rules, namely R→ L / _H and R→ L / __F (that is, *213 → 21 / _45, and *213 → 21 / _53) neither of which is actually attested. Another issue is that there is no evidence that these primitive level tones, required by this analysis, are active elsewhere in the phonology.

It is worth reiterating that sandhi in Tianjin is motivated and constrained via feature specifications in the Tianjin feature hierarchy: as with Beijing Mandarin and Pingyao, phonological processes that involve changes in value of single features are preferred, and we would not normally expect, for example, fully specified tones that do not share any features with each other, i.e. /53/ and /213/ on the one hand, and /21/ and /45/ on the other, to alternate in a symmetrical system.

Consider the following three-syllable sandhi processes that have been reported for Tianjin, all of which conform to the prediction that single feature changes are preferred. Constituent

---

[42] Although this might imply /35/ would better reflect phonetic reality of the tone than /45/, we utilize the latter here in keeping with established scholarship.

structure is given in square brackets. Chen (2000:109) argues that the application of trisyllabic sandhi is blind to constituent structure in Tianjin:

(46) Tianjin three-syllable sandhi as reported by Chen (2000),[43] adapted

| | | | | | |
|---|---|---|---|---|---|
| i. | /53/ | /53/ | /21/ | → | **21** **45** 21 |
| | [*si* | *ji]* | *qing* | | ‘evergreen’ |
| | *zuo* | *[dian* | *che]* | | ‘take a train’ |
| ii. | /213/ | /213/ | /213/ | → | **45** **45** 213 |
| | [*li* | *fa]* | *suo* | | ‘barber shop’ |
| | *mu* | *[lao* | *hu]* | | ‘tigress’ |
| iii. | /53/ | /53/ | /53/ | → | **45** **21** 53 |
| | [*su* | *liao]* | *bu* | | ‘plastic cloth’ |
| | *ya* | *[re* | *dai]* | | ‘subtropical’ |
| iv. | /21/ | /21/ | /21/ | → | 21 **213** 21 |
| | [*tuo* | *la]* | *ji* | | ‘tractor’ |
| | *kai* | *[fei* | *ji]* | | ‘pilot a plane’ |
| v. | /213/ | /21/ | /21/ | → | **45** **213** 21 |
| | [*bao* | *wen]* | *bei* | | ‘thermos’ |
| | *da* | *[guan* | *qiang]* | | ‘speak in a bureaucratic tone’ |
| vi. | /21/ | /53/ | /53/ | → | **213** **21** 53 |
| | [*wen* | *du]* | *ji* | | ‘thermometer’ |
| | *tong* | *[dian* | *hua]* | | ‘make a phone call’ |

43 Glosses for individual words are not provided for this data in Chen (2000).

| | | | | | |
|---|---|---|---|---|---|
| vii. | /53/ | /21/ | /21/ | → | 53 **213** 21 |
| | [*lu* | *yin]* | *ji* | | 'cassette recorder' |
| | *shang* | *[fei* | *ji]* | | 'board an airplane' |

There are additional trisyllabic sandhi processes. Chen (2000:106) writes that of the 64 possible combinatorial possibilities (i.e. $4^3$), 37 are sandhi free, twenty call for only a single application of a sandhi rule (e.g. /53 21 213/ -> [**45** 21 213] via (31) above), and the remaining "seven tonal combinations potentially involve more than one sandhi process" (2000:107). In order to arrive at the respective outputs, two-syllable sandhi rules are applied in succession, sometimes creating new sandhi contexts over the course of the derivation, eventually yielding the output form, as below. The domain of application is indicated with square brackets, thus:

(47) Trisyllabic sandhi, pattern *v*

| | | | | |
|---|---|---|---|---|
| 213.[21.21] | → | [213.**213**].21 | → | **45.**213.21 |
| | | via (30) | | via (29) |

The initial application of the first bisyllabic sandhi process on the rightmost sandhi window thus creates the context for the second application of a sandhi process on the leftmost window. One might immediately notice that the domain of application in (47) begins with the rightmost two-syllable window, and progresses to the leftmost two-syllable window. This in mind, consider the following cases in which only a single derivational path is correct despite there being multiple possible paths with different combinations of sandhi application:

(48) Trisyllabic sandhi, patterns *iii*, *iv*, *vi*, and *vii*

a. pattern *iii, left to right*

| | | | | |
|---|---|---|---|---|
| [53.53].53 | → | **21**.[53.53] | → | *21.**21**.53 |
| | | via (32) | | via (32) |

b. pattern *iii*, *right to left*

| | | | | |
|---|---|---|---|---|
| 53.[53.53] | → | [53.**21**].53 | → | **45.**21.53 |
| | | via (32) | | via (31) |

---

*c. pattern iv, left to right*

| | | | | |
|---|---|---|---|---|
| [21.21].21 | → | **213**.[21.21] | → | *213**.213**.21 |
| | | via (30) | | via (30) |

*d. pattern iv, right to left*

| | | | | |
|---|---|---|---|---|
| 21.[21.21] | → | 21.**213**.21 | | |
| | | via (30) | | |

---

*e*. pattern *vi, left to right*

| | | | | |
|---|---|---|---|---|
| [21.53].53 | → | 21.[53**.**53] | → | *21.**21.**53 |
| | | | | via (32) |

*f*. pattern *vi, right to left*

| | | | | |
|---|---|---|---|---|
| 21.[53.53] | → | [21.**21**].53 | → | **213**.21.53 |
| | | via (32) | | via (30) |

---

g. *pattern vii, left to right*

| | | | | |
|---|---|---|---|---|
| [53.21].21 | → | **45**.[21.21] | → | *45**.213**.21 |
| | | via (31) | | via (30) |

*h. pattern vii, right to left*

| | | | | |
|---|---|---|---|---|
| 53.[21.21] | → | [53.**213**].21 | | |
| | | via (30) | | |

---

Again, in all cases here, although multiple derivational pathways are possible, the only derivations that yield the desired results proceed from right to left. The default assumption would therefore be that trisyllabic sandhi in Tianjin proceeds from the rightmost two-syllable window, and moves incrementally to the left. This, however, turns out not to work for other reported patterns:

(49) Trisyllabic sandhi, patterns *i* and *ii*

| | | | | |
|---|---|---|---|---|
| a. *pattern i, right to left* | | | | |
| 53.[53.21] | → | *53.**45.**21 | | |
| | | via (31) | | |
| b. *pattern i, left to right* | | | | |
| [53.53].21 | → | **21**.[53.21] | → | 21.**45**.21 |
| | | via (32) | | via (31) |
| c. *pattern ii, right to left* | | | | |
| 213.[213.213] | → | *213.**45.**213 | | |
| | | via (29) | | |
| d. *pattern ii, left to right* | | | | |
| [213.213].213 | → | **45**.[213.213] | → | 45**.45**.213 |
| | | via (29) | | via (29) |

As can be seen, some of the derivation pathways proceed from right to left, and others from left to right in order to arrive at the desired output. A number of attempts (Chen 1986, 2000; Hung

1987; Tan 1987; Zhang 1987) from various frameworks have been made to explain this apparent ‘paradox’ (Chen 1986, 2000) of rule ordering. Indeed to date there appears to be ‘no consensus’ (Li and Chen 2012) reached. A possible explanation for this is provided in Li and Chen (2012), who state that it is worth noting that while all of the studies on trisyllabic sequences in Tianjin Mandarin have been based on impressionistic observations, more recent empirical studies show that one of the actual bisyllabic sandhi rules may be different from the descriptions oft-repeated in the literature, implying that the trisyllabic patterns in Tianjin may also differ from those described in the literature. Li and Chen (2012) report that /53 53/ sequences do not generally result in sandhi, and Zhang and Liu (2011), in agreement with findings in Shi (1988), Shi and Wang (2004), Liu and Gao (2003) and Gao (2004) report that the rule is applied inconsistently, if at all, and, when applied does not actually result in an actual [**21.**53] output sequence, arguing that “the ‘sandhi’ tone shares much resemblance to the /53/ tone in the second syllable and is considerably higher in pitch than /21/, the tone it supposedly neutralizes to” (Zhang and Liu 2011:26). Li and Chen (2012) provide empirical data for all of the reported bisyllabic and trisyllabic sandhi sequences in Tianjin, and argue that there is not any compelling evidence for the /53/ → 21 / __ /53/ rule in Tianjin. Taken alongside evidence by Zhang and Liu (2011, 2016), we might hypothesize that it was a historical process (possibly with a number of exceptions) that has since fallen out of use (the structure of Proto-Tianjin would have permitted the existence of this process before reanalysis).

Assuming then that /53/ → 21 / __/53/ is not an actual sandhi process, we can vastly simplify tone sandhi in Tianjin, as below (for the sake of brevity I omit trisyllabic sandhi involving the computation of only a single process, e.g. /213 213 21/ → [**45** 213 21]):

| | | | | |
|---|---|---|---|---|
| | *'Tone 3 sandhi'* | 213 → 45 / __213 | | |
| Bisyllabic | *Falling tone dissimilation a* | 21 → 213 **/** __21 | | |
| sandhi | *Falling tone dissimilation b* | 53 → 45 / __21 | | |
| Trisyllabic | *pattern i* | /213 213 213/ | → | **45** **45** 213 |
| sandhi | *pattern ii* | /21 21 21/ | → | 21 **213** 21 |
| | *pattern iii* | /213 21 21/ | → | **45** **213** 21 |
| | *pattern iv* | /53 21 21/ | → | 53 **213** 21 |

Table 8: Tianjin sandhi, revised

The somewhat simple explanation for the 'Tianjin paradox' thus is that there are two types of sandhi in Tianjin with different origins: falling tone dissimilation (specific to Tianjin), and Tone 3 sandhi (shared with most other Mandarin languages). Trisyllabic sandhi in Tianjin proceeds from the right edge for falling tone dissimilation, after which Tone 3 sandhi is applied from left to right:

(50) Tianjin revised trisyllabic sandhi derivation

*pattern i*

[213.213].213 → **45**.[213.213] → **45.45.**213

via '*Tone 3 sandhi*' (left to right)

n.b. 213.213.213 → *****213.**45**.213

*pattern ii*

21.[21.21] → 21.**213.**21

via '*falling tone dissimilation a*' (right to left),

n.b. 21.21.21 → ***213.213**.21

*pattern iii*

213.[21.21] → [213.**213**].21 → **45.213.**21

via '*falling tone dissimilation a'* (right to left)

via '*Tone 3 sandhi*' (left to right)

*pattern iv*

53.[21.21] → 53.**213.**21

via '*falling tone dissimilation a*' (right to left)

n.b. 53.21.21 → ***45.**21.21 → ***45.**213**.21

Interestingly, the simplified picture concords with our insights above, afforded by the contrastive framework, that Tianjin underwent a reanalysis. The contrastive reanalysis for Tianjin, postulated above, was enabled by the high degree of phonetic drift of tones /21/ and /45/, the modern correlates of MC tones Ia and Ib, previously phonetically high and low respectively. Given the relative similarity in the modern forms of MC tones II and III in both Beijing (/214/ and /51/ respectively) and Tianjin (/213/ and /53/ respectively) we can postulate minimal drift in these tones since bifurcation. That the low falling tone in Tianjin, /21/, was once higher however implies the likelihood that at some time in the past it was phonetically much closer to the high falling tone, /53/. This possibility is bolstered by the fact that all of the sandhi sequences acquired after Tianjin split from the Beijing Mandarin branch involve sequences of two falling tones: since all of the sandhi processes unique to Tianjin are dissimilatory in nature and involve sequences of two falling tones, we might speculate that they may have been motivated by the historical need to distinguish the two phonetically similar falling tones in sequence.  Regardless of motivation however, it is noteworthy that the two falling tone dissimilatory processes which have their origins in the modern Tianjin hierarchical configuration are derived in sandhi contexts first and proceed from right to left, while the 'Tone 3 sandhi' process, which originated in a different contrastive system entirely is applied last in the derivation[44] and proceeds from left to right.[45]

### 3.2.5 **Yantai Mandarin**

Yantai Mandarin is a member of the Jiao-liao family of dialects spoken on the Shangdong peninsula. Like Pingyao, Yantai possesses four citation tones, with two tones being merged in citation forms, but exhibiting differential behaviour in sandhi contexts. Also like Pingyao, Yantai exhibits a fifth tone that does not exist independently and serves only as an output of sandhi processes:

(51) Yantai Mandarin tone inventory (adapted from Bao 1999; Chen 2000)

| | | | |
|---|---|---|---|
| i | *fu* | /31a/ | 'man' |
| ii. | *ren* | /31b/ | 'person' |
| ii | *fa* | /214/ | 'method' |
| iii | *tʰu* | /55/ | 'picture' |
| iv. | | [35] | |

44 Interestingly, the same process in Pingyao, /II.II/ → /Ib.II/ is also always applied last in the derivation in that language as well.

45 This may superficially appear to be a counterexample to the tendency, described by Bermúdez-Otero (2007, 2015), for older processes to occur in increasingly narrow cyclic domains, which precede wider ones in feed-forward derivations. However there is no counterevidence here to the general unidirectionality of this life cycle for any individual process. Bermúdez-Otero (2015:21) argues that "the speed with which individual phonetic and phonological processes rise through the grammar is contingent" on a variety of factors, and we can account for this simply by postulating independent evolutionary trajectories for each of the rules.

Tone /31a/ can be differentiated from /31b/ by observing the following two sandhi rules:

(52) Yantai Mandarin sandhi rule 1 (Bao 1999, Huang 2012)[46]

/31a/ → 35 / __ /31/

| | | | | | |
|---|---|---|---|---|---|
| i. | *san* | *p*ʰ*o* | | *san* | *p*ʰ*o* |
| | | | → | ‘hill slope’ | |
| | /31a/ | /31/ | | **35** | 31 |
| ii. | *tɕy* | *tsung* | | *tɕy* | *tsung* |
| | | | → | ‘pig's bristles’ | |
| | /31a/ | /31/ | | **35** | 31 |

(53) Yantai Mandarin sandhi rule 2 (Chen 2000, Huang 2012)

/31b/ → 55 / __ /31/

| | | | | | |
|---|---|---|---|---|---|
| i. | *ja* | *kao* | | *ja* | *kao* |
| | | | → | **‘**toothpaste’ | |
| | /31b/ | /31/ | | **55** | 31 |
| ii. | *mian* | *xua* | | *mian* | *xua* |
| | | | → | ‘cotton’ | |
| | /31b/ | /31/ | | **55** | 31 |

46 Data is from Bao (1999); Chen (2000); Huang (2012). Glosses for individual words is not provided in the original source.

The following two dissimilatory sandhi rules in Yantai involve the change of one citation tone into another.

(54) Yantai Mandarin sandhi Rule 3 (Bao 1999, Huang 2012)

/214/ → 55 / __ /214/

| | | | | | |
|---|---|---|---|---|---|
| i. | *ɕiao* | *mo* | | *ɕiao* | *mo* |
| | | | → | 'wheat' | |
| | /214/ | /214/ | | **55** | 214 |
| ii. | *y* | *sui* | | *y* | *sui* |
| | | | → | 'rain water' | |
| | /214/ | /214/ | | **55** | 214 |

(55) Yantai Mandarin sandhi Rule 4 (Bao 1999, Huang 2012)

/55/ → 31 / __ /55/

| | | | | | |
|---|---|---|---|---|---|
| i. | *ɕy* | *pʰi* | | *ɕy* | *pʰi* |
| | | | → | 'tree bark' | |
| | /55/ | /55/ | | **31** | 55 |
| ii. | *tɕiang* *ɕi* | | | *tɕiang* | *ɕi* |
| | | | → | 'to stage (a) performance' | |
| | /55/ | /55/ | | **31** | 55 |

The remaining sandhi rule, below, involves the change of the concave tone into the non-citation tone [35].

(56) Yantai Mandarin sandhi Rule 5 (Bao 1999, Huang 2012)

/214/ → 35 / __ /31/

| | | | | | |
|---|---|---|---|---|---|
| i. | *ɕiu* | *ɕin* | | *ɕiu* | *ɕin* |
| | | | → | 'palm' | |
| | /214/ | /31/ | | **35** | 31 |

| | | | | |
|---|---|---|---|---|
| ii. | *ɕiao* | *san* | | *ɕiao* *san* |
| | | | → | ‘small hill’ |
| | /214/ | /31/ | | **35** 31 |

Phonological activity in Yantai can be summarized in the table below:

| Input | Outputs | Rule |
|---|---|---|
| /31a/ | 35 | (52) |
| /31b/ | 55 | (53) |
| /214/ | 55, 35 | (54), (56) |
| /55/ | 31 | (55) |

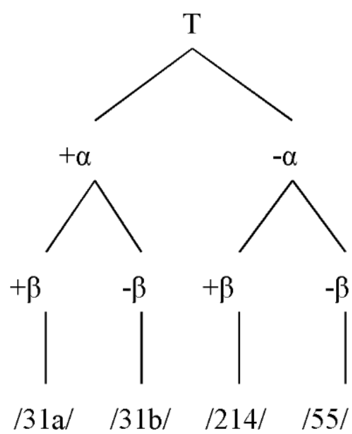
Table 9: Yantai sandhi alternations.

Observing the table above, we can deduce that /31b/ and /55/ differ by a single feature, that /55/ and /214/ differ by a single feature and therefore, by implication /31b/ and /214/ share no features, and, assuming a two-feature tree, /31a/ and /31b/ differ by a single feature, and /31a/ and /214/ also differ by a single feature. It is also reasonable to assume that either of /31a/ or /214/ are allotonic with the non-citation tone [35] as it is the target of processes involving only these as inputs, confirming our supposition that /214/ and /31a/ also differ by a single feature (since an allotonic variant of one of them is a target of the other).[47] We also must contend with the fact that rule 4 in (55) involves either of /31a/ or /31b/ as outputs to a rule involving /55/ as an input, but

[47] Although there is no conclusive evidence concerning which of /214/ and /31a/ possesses the allotonic variant [35], we might suspect /31a/ as a more likely candidate if we consider the patterning of Pingyao.

since the two are merged in citation form, the question as to which is the target is undecidable without knowing the full feature representation of each tone . This yields the following tree:

(57) Yantai Mandarin feature hierarchy (abstract features):[48]




Unlike our previous three languages, the tonal categories of Yantai do not correspond neatly to those of Beijing, Pingyao and Tianjin, as can be seen in Table 10 below. Although the tonal categories are identical in all other respects, /55/ in Yantai, which broadly corresponds to MC tone III, also possesses a group of words that historically belonged in the MC tone Ib category, unlike the other three languages. This implies that Yantai, at some point in the past, experienced a partial merger between Ib and III that did not occur in other languages. This is not an outcome we would expect if Yantai possessed the same featural specifications of the same modern correlates of MC tonal categories as Mandarin and Pingyao, since the two categories share no features.

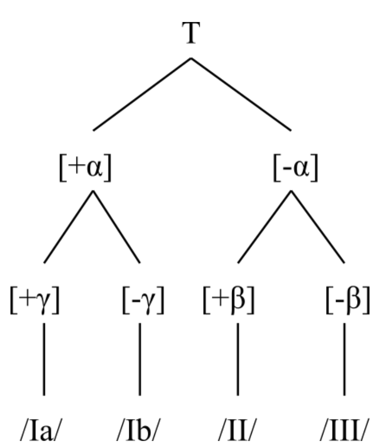
48 One possible phonetic labelling of these features would be [+rising] for [+β] (assuming [35] as a rising allotone of 31a) and [-rising] for [-β], [+falling] for [+α], and [-falling] for [-α].

| MC Tone | MC cognate | Beijing | Yantai | Character | English gloss |
|---|---|---|---|---|---|
| | | /55/ [+α +β] | /31a/ [+α +β] | | |
| Ia | fio | fu | fu | 夫 | ‘man/husband’ |
| | tɑu | dao | tao | 刀 | ‘knife’ |
| | $t^h$ɑu | tao | $t^h$aɔ | 滔 | ‘overflow’ |
| | | /35/ [+α -β] | /31b/ [+α -β] | | |
| Ib | ŋa | ya | ja | 牙 | ‘tooth’ |
| | mian | mian | mian | 棉 | ‘cotton’ |
| | | | **/55/** | | |
| | do | tu | $t^h$u | 图 | ‘diagram’ |
| | dɑu | tao | $t^h$ɔ | 桃 | ‘peach’ |
| | | /214/ [-α -β] | /214/ [-α +β] | | |
| | tɑu | dao | tao | 岛 | ‘island’ |
| | $t^h$ɑu | tao | $t^h$ao | 讨 | ‘to ask for’ |
| II | | | | | |
| | | /51/ [-α +β] | /55/ | | |
| | dɑu | dao | tao | 道 | ‘road’ |
| | | /51/ [-α +β] | /55/ [-α -β] | | |
| III | dɑu | dao | tao | 盗 | ‘rob’ |
| | tɑu | dao | tao | 到 | ‘arrive’ |

Table 10: Beijing and Yantai tonal categories and Middle Chinese cognates

Yantai does not possess the same featural specifications, however, and at some point, categories II and III, which are contrastive sisters, swapped their respective [β] values as below:

(58) Middle Chinese tonal cognates of Yantai:



At this point, in the absence of additional evidence regarding the phonetic shape of Middle Chinese tones we have little to say regarding specific motivation for the tone Ib -> III merger, save that it was enabled by a contrastive shift of [β] values in the [-α] domain. This does, however, reinforce the notion that evidence from activity patterns has the potential to shed significant light on diachrony.

### 3.2.6 **Nantong Mandarin**

Nantong is quite unique amongst the Mandarin languages. It is mutually unintelligible with all other dialects save for the closely related Jinsha Mandarin. Ao (1993) argues that the uniqueness of Nantong is deeply associated with its history. Until the 3rd century AD, the Nantong area in the Yangtse river delta was completely submerged under water. By the end of that century a large sandbar island, *Hudouzhou*, emerged in the middle of the river, remaining as an island until the ninth century when the northern waterway silted up and became part of the north bank. The *Taiping Huanyu Ji* [The Peaceful World], a geographical and historical compendium published in the tenth century, describes the inhabitants of *Hudouzhou* in the mid-6th century as exiles who made salt from sea water for a living. The location of *Hudouzhou* makes it likely that the original inhabitants came from neighbouring areas and spoke a variety of dialects, with their speech eventually merging into a single variant with mixed characteristics (Ao 1993). Of relevance to the present study is the fact that Nantong can be regarded as chiefly Mandarin, with some aspects more characteristic of Wu family languages (more than four tones, the bifurcation of MC TIII, and the existence of a ‘neutral tone’). This is unsurprising as the Yangtse river is a major isogloss with regions directly south largely dominated by Wu languages, and regions to the north largely dominated by Mandarin languages (Chao 1976; Norman 1988).

Nantong Mandarin possesses the following inventory of citation tones. Two of these tones occur only on obstruent-final syllables and are short, or checked tones, *vi* and *vii*, and can be considered allotonic to their full tone counterparts. Some morphemes, typically with minimal semantic content (e.g. past tense markers, third person pronouns) are underlyingly toneless (Ao 1993:106), with pitch height being predictable by environment. We notationally indicate this in *viii* as a ‘neutral’ tone, [33].

(59) Nantong Mandarin inventory (adapted from Ao 1993:63-4)[49]

| | | | |
|---|---|---|---|
| i. | *ti* | /31/ | ‘ladder’ |
| ii. | *ti* | /44/ | ‘body’ |
| iii. | *ti* | /212/ | ‘ground’ |
| iv. | *ti* | /24/ | ‘raise’ |
| v. | *ti* | /51/ | ‘go’ |
| vi. | *tiʔ* | [51q] | ‘iron’ |
| vii. | *tiʔ* | [44q] | ‘flute’ |
| viii. | | [33] | |

[49] Ao (1993) utilizes high/mid/low primitives to designate tone levels, e.g. /24/ above is MH and /44/ is H. The Chao tone numerals used above were arrived at by utilizing raw average F0 values (Ao 1993:64) for start/mid/endpoints of the tones, determining the number of half-steps between these values, and then scaling to a 5-level distinction to yield roughly equal perceptual space between each numeral.

Ao (1993) argues that Nantong has three major types of tonal alternations: tone spreading, tone deletion, and regressive tone change. Of these, it seems likely that only regressive tone change is a true form of tone sandhi involving feature change.

Tone spreading is the process which determines the ultimate phonetic form that an underlyingly toneless syllable is realized as in a given tonal context. When a toneless syllable occurs word-initially, it is always realized as a neutral, mid tone, [33]. In word-medial environments toneless syllables are realized as either high or mid, depending only on the tone of the preceding syllable, with /24/ in initial position conditioning a high tone realization on the following syllable, and /212/, /51/, and /31/ conditioning a mid tone realization on the following syllable. /44/ conditions either, and the realization is argued to be optional. In cases where toneless syllables occur word finally, they may be realized as either low, mid or high, depending upon the preceding tone, and strings of multiple toneless syllables to the right of a toned syllable are realized at the same height, as below:

| Underlying Initial Tone | 1st toneless σ | Subsequent toneless σ |
|---|---|---|
| /31/ | L | L |
| /44/ | L | L |
| | H | H |
| /212/ | M | M |
| /24/ | H | H |
| /51/ | L | L |

Table 11: Nantong Syllable final tonal spreading. (adapted from Ao 1993:106)

Accompanying this process is a form of tonal redistribution (Ao 1993:108), which is the disassociation of a tone from one syllable and reassociation with another in some environments where toneless syllables are present. When /212/ or /24/ precede a toneless syllable, the full contour of the tone is spread over two syllables, as below:

(60) Nantong tonal redistributrion[50]

| | | | | |
|---|---|---|---|---|
| i. | *ŋo* | *ɚ* | → | *ŋo . ɚ* 'child' |
| | /24/ | Ø | | [2 . 4] |
| ii. | *jə* | *tə* | → | *jə . ɾə* 'pardoned' |
| | /24/ | Ø | | [2 . 4] |
| iii | *jə* | *tə* | → | *jə . ɾə* 'circled' |
| | /212/ | Ø | | [21 . 2] |
| iv. | *poŋ* | *ɚ* | → | *poŋ . ɚ* 'stick' |
| | /212/ | Ø | | [21 . 2] |

50 All data is from Ao (1993). Glosses for individual words are not provided in the source.

Neither tonal redistribution nor tone spreading take toned syllables as targets and there is no evidence that these processes involve feature change. Ao makes the observation (1993:110) that toneless syllables which are the target of tone spreading also usually undergo some degree of onset lenition, suggesting that these processes are possibly simply a byproduct of connected speech.

Tone deletion is another mechanical process that Ao argues occurs in three-syllable words consisting of three toned syllables. In connected speech, the second syllable of a trisyllabic word "loses its underlying tone and may become the target of tone spreading" (Ao 1993:114). It is always the second syllable which loses its tone and in doing so also becomes a site of onset lenition.

Nantong also has several robust sandhi processes. Interestingly, the sandhi context in Nantong often includes not only the following tone, as is the case with most Mandarin languages, but also whether the target is word-initial or word-medial, as described below. While the low concave tone, /212/, is realized as [24] before a falling (/31/, /51/) or second concave tone (/212/) and as [33] (a non-citation tone) before a high (/44/) or rising (/24/) tone when in word-initial position, it is realized instead as [31] before a rising, high, or high-falling tone (/24/, /44/, /51/) and as [24] before a mid-falling or concave tone (/31/, /212/) in word-medial position:

(61) Nantong concave tone sandhi (adapted from Ao 1993)

(a) Word-initial I

| | | | | |
|---|---|---|---|---|
| | /212/ → | 24 / [ __{/212/, /31/, /51/} ]$_\omega$ | | |
| *i.* | *tiŋ* *ti* | | *tiŋ* | *ti* |
| | | → | | |
| | electric ladder | | 'elevator' | |
| | /212/ /31/ | | **24** | 31 |

ii. *tiŋ* xʷo → *tiŋ* xʷo

electric speech → 'telephone'

/212/ /212/ → **24** 212

(b) Word-initial II

/212/ → 33 / [ __{/44/, /24/, /4q/}]$_\omega$

i. *tiŋ* *jɨŋ* → *tiŋ* *jɨŋ*

electric image → 'movie'

/212/ /44/ → 33 44

ii. *tiŋ* *lβ̩* → *tiŋ* *lβ̩*

electric stove → 'electric stove'

/212/ /24/ → 33 24

(c) Word-medial I

/212/ → 31 / [T __{/44/, /24/, /51/} ]$_\omega$, T = any tone

i. *tiŋ* xʷo *woŋ* → *tiŋ* xʷo wo*ŋ*

electric speech net → 'telephone network'

/212/ /212/ /44/ → 212 **31** 44

ii. *tiŋ* xʷo ɕi*ŋ* → *tiŋ* xʷo ɕi*ŋ*

electric speech line → 'telephone line'

/212/ /212/ /51/ → 212 **31** 51

(d) Word-medial II

/212/ → 24 / [T __ {/31/, /212/}]$_{\omega}$, T = any tone

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *i.* | *tiŋ* | xʷo | tʃɜ[51] | | *tiŋ* | xʷo | tʃɜ |
| | | | | → | | | |
| | electric speech net | | | | 'telephone network' | | |
| | /212/ | /212/ | /31/ | | 212 | **24** | 31 |

The low- or mid-rising tone /24/ is also realized as the non-citation tone [33] before a high flat tone in word-initial position, as in (62) below:

(62) Nantong Low-rising tone sandhi (adapted from Ao 1993)

/24/ → 33 / [ __ /44/ ]$_{\omega}$

| | | | | | |
|---|---|---|---|---|---|
| | *xu* | ɕʷe | | *xu* | ɕʷe |
| | | | → | | |
| | river | water | | 'river water' | |
| | /24/ | /44/ | | 33 | 44 |
| | | | | | |
| *i.* | *xu* | ɕʷe | | *xu* | ɕʷe |
| | | | → | | |
| | river | water | | 'river water' | |
| | /24/ | /44/ | | 33 | 44 |

Finally, the mid-falling tone, /31/ is also realized as the non-citation tone, [33] before **any** tone, when in word-initial position. This does not occur in word-medial position, where it retains its citation form:

---

51 *tʃɜ* here and *woŋ* in (c) are distinct words despite having the same English gloss. The character for *woŋ* is 网 while *tʃɜ* likely corresponds to 净**.**

(63) Word Initial 4

/31/ → 33 / [ __ T]$_{\omega}$ , T = any tone

| | | | | | |
|---|---|---|---|---|---|
| *i.* | *fe* | *tiŋ* | → | *fe* | *tiŋ* |
| | flying | boat | | 'blimp' | |
| | /31/ | /44/ | | 33 | 44 |
| *ii.* | *fe* | *tɔŋ* | → | *fe* | *tɔŋ* |
| | flying | bullet | | 'missile' | |
| | /31/ | /212/ | | 33 | 44 |

We can summarize the alternations for Nantong as follows:

| Input | Outputs | Rule |
|---|---|---|
| /212/ | 24, 31, 33 | (61a-d) |
| /24/ | 33 | (62) |
| /31/ | 33 | (63) |

Table 12: Nantong sandhi tonal alternations.

We thus deduce that /212/ differs by a single feature from both /24/ and /31/ and either underlies or differs by a single feature from whichever tone underlies the allotonic variant, [33]. Similarly /24/ and /31/ both either differ by a single feature from, or themselves underlie the allotonic variant [33]. To sort this issue out, let us first look at the modern tone correlates of Middle Chinese tonal categories proposed by Ao:

| MC Onset | MC TI | MC TII | MC TIII | MC TIV |
|---|---|---|---|---|
| voiceless | /31/ (Ia) | **/44/** | /51/ | /51q/ |
| voiced | /24/ (Ib) | **/212/** | /212/ | /44q/ |

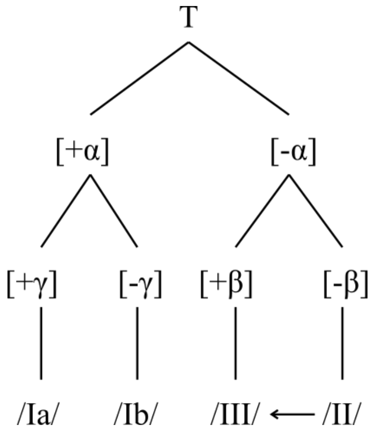
Table 13: Nantong tones by Middle Chinese categories (adapted from Ao 1993:65)

A number of characteristics distinguish Nantong tonally from the other Mandarin languages investigated thus far. As seen in the table above, unlike most Mandarin languages, Nantong preserves a class of closed syllables with checked tones allotonic to their full-tone counterparts that were lost in most other Mandarin languages.[52] Nantong also exhibits a trace of the dissimilatory sandhi targeting the first of two adjacent modern correlates of the Middle Chinese tone II common in other Mandarin languages (/212/ → 24 / __/212/, in Example (61a.ii) above); however, there are a few notable differences. In addition to sandhi being triggered by a second low concave tone, it is also triggered by the high-falling tone, /51/, and the mid-falling tone, /31/. Also, and of most relevance to the present discussion, this sandhi only occurs on MC tone II words with formerly voiced initials. The contrast between formerly voiced and voiceless initials in Middle Chinese, reanalyzed as a tonal contrast in the tone I category, was lost in the MC tone II and III categories in other Mandarin languages. In Nantong however, there is evidence that tone III also bifurcated into high and low variants based on syllable-initial voicing contrasts. Table 13 indicates that, where there was in Middle Chinese a four-way contrast (tone II voiced initial, tone II voiceless initial, tone III voiced initial, tone III voiceless initial), there is now a three-way contrast in modern Nantong (tone II voiced initial, tone III voiced initial, tone II/III voiceless initial). This implies one of two possibilities: either both tones II and III bifurcated before the low variants merged again, or a partial merger occurred between the two tonal categories, followed by a split in one of them. I assume the latter case here, as tones II and III were contrastive sisters in late Middle Chinese and minimality considerations suggest partial mergers between contrastive sisters to be the most likely

52 Pingyao shares this characteristic with Nantong.

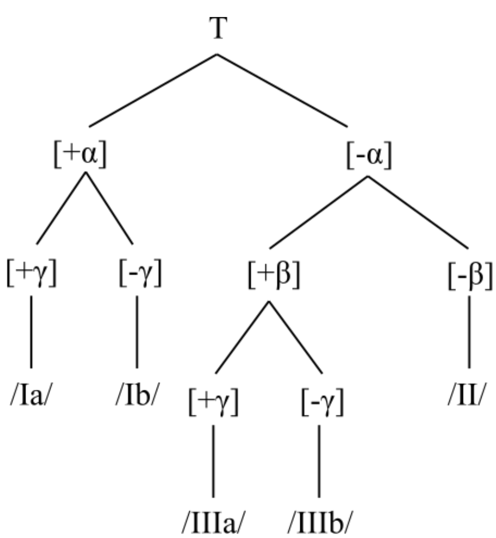
(if both tones bifurcated prior to merging, we need to postulate a higher-level merge which is less parsimonious). In addition, we already know that there was a partial merger in Middle Chinese of tone II into tone III, discussed in more depth in §3.2.7, below. We thus start with the contrastive specifications of Middle Chinese, at the time of our partial Tone II → III merger:

(64) Middle Chinese / Proto-Nantong Stage 1, Tone II → III merge.



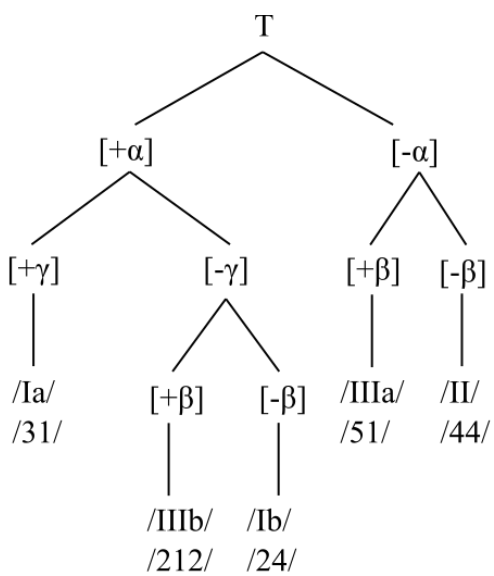

From here we postulate a bifurcation of tone III at the most subordinate level on the basis of height, [γ], as initial voicing contrasts were lost:

(65) Proto-Nantong Stage 2, Tone III split



This tree was later restructured, evidenced by the fact that tones IIIb (/212/) and Ia (/31/), which share no features in Proto-Nantong Stage 2, and IIIb (/212/) and Ib (/24/), which differ by two features in Proto-Nantong Stage 2, undergo sandhi alternations in modern Nantong. In addition, each of IIIb, Ia, and Ib undergo an alternation with the allotonic variant [33], but no other tones do, suggesting that each of these three must differ by a single feature from each other. We thus postulate a reanalysis whereby IIIb → [+α] ([-α +β - γ] → [**+α** +β -γ]), introducing a [β] contrast at the most subordinate level of the formerly terminal Ib node, and rendering the [γ] feature of IIIa redundant, as below:

(66) Modern Nantong




It is worth noting that with Nantong, as in Tianjin, the reanalysis of /IIIb/ as [+α] is motivated on phonetic grounds, with all other [-α] tones being contrastively [+high] and [+α] tones being contrastively [-high]. This is supported by the observation that [γ] was previously a height contrast, and /Ia/ was once contrastively high, subsequently drifting downwards, and enabling the repurposing of the [γ] feature as a contour contrast.

Feature labelling is somewhat less straightforward in Nantong. A brief description is given here, and discussed in depth later on. [α] denotes a height contrast. Positive values of [β] are falling whereas negative values are not falling. [γ] denotes a second contour contrast. Positive values are [falling] and negative values are rising. The concave tone, /212/, is thus conceptualized as the application of two functions, [falling] and [rising], and the specifics of how this is realized is taken up in §5.4.1.

In the settling of the question of what the underlying tonal identity of the allotonic variant [33] is, we find further evidence for the structure above. Based on the tree structure above, we should immediately identify /31/ as the most likely candidate as the underling form of [33], since the set of alternations involving [33] then only involves single feature changes (i.e. /212/ → [+γ] and /24/ → [+γ]), whereas multiple feature changes need to be invoked for /31/ if either of /212/ or /24/ is the underlying tonal identity for [33]. This in mind, the unusually wide distribution of the sandhi form exemplified by (63) above (/31/ → 33 / [ __ T]$_{\omega}$, T = any tone) is accounted for best if [33] is an allotonic variant of /31/ rather than any other tone.

Another possibility is that [33] is archiphonemic (see Spahr 2014). Spahr argues that segments in neutralized positions can be represented as non-terminal nodes of the contrastive hierarchy. This is accomplished by deleting feature values reflecting the contrast which is lost from the terminal nodes upwards. Neutralized forms thus do not contain the full representation of either member of a neutralized pair, but a third archiphonemic segment. Spahr (2014:552) makes the argument that this approach is conceptually preferable because it does not require the specification of contrastive features in neutralized contexts, and because neutralized segments in general often exhibit intermediate phonetic realizations relative to their corresponding non-neutralized counterparts, which is expected if the contrastive feature defining the boundary between the two segments is deleted. Assuming this type of analysis, in the case of Nantong [33] is specified simply as [+α], ([-high]).[53]

### 3.2.7 Diacronic processes and a Middle Chinese hierarchy

The primary focus of this section is concerned with utilizing diachronic evidence to further supplement our understanding of the respective feature hierarchies for each language in the previous section. Thus we will start with a brief review of the various relevant phonological changes that have been argued to occur between Middle Chinese and modern Mandarin, and later discuss the ordering of these general processes as well as various language-specific processes in each of the above languages, and discuss what implications this has for our own hierarchical representations.

---

53 This analysis works well for Nantong, however the phonetic basis for an archiphonemic representation of the non-citation tones [31] in Pingyao and [35] in Yantai, both as [+α] is less clear. This question is left to future research.

### 3.2.7.1 **Middle Chinese to Modern Mandarin**

Middle Chinese, from which all languages in the present study originate, was initially a three-, and later a four-tone system. Although the precise phonetic details of this system have yet to be fully articulated, Chen (1976; also see Pulleyblank 1978) argues that it is still possible to state the tonal changes that transpired between Middle Chinese and modern Beijing Mandarin as mergers or splits between categories, and postulates three tonal changes of relevance: a register split of tone I, a partial merger of tone II with tone III, and a redistribution of tone IV (the 'checked' tone in MC occurring only with closed syllables) amongst all other categories. Interestingly, all three of these processes seem to be conditioned by initial voicing, as below:

(67) Tonal changes from MC to Proto-Mandarin/Jin.

(a) Tone I split

$$\text{T1} \quad \rightarrow \quad \begin{cases} \text{T1a / voiceless initial} \\ \text{T1b / voiced initial} \end{cases}$$

(b) Tone II merge

$$\text{T2} \quad \rightarrow \quad \text{T3 / voiced obstruent initial}$$

(c) Tone IV tripart

$$\text{T4} \quad \rightarrow \quad \begin{cases} \text{T3 / sonorant initial} \\ \text{T1b / voiced obstruent initial} \\ \text{T1a, T1b, T2, T3 / voiceless initial} \end{cases}$$

The voiced/voiceless contrast of syllable-initial obstruents that was later subject to neutralization by devoicing is thus preserved through the Tone I split as a tonal rather than a segmental contrast. Similarly, Middle Chinese Tone II devoiced initial obstruents were reanalyzed as contrastively Tone III. MC Tone IV was a somewhat restricted category, occurring only in checked syllables. While the Chinese philological tradition traditionally treats Tone IV in Middle Chinese as a separate category, contemporary scholarship treats it as a non-contrastive allotone of one of the other three MC tones (Baxter 1992; Chen 2000). Chen (1976) argues that, as the stop endings were lost, the checked syllables merged with the open ones and obligatorily assumed a contrastive pitch contour. If it is in fact the case that the loss of stop endings triggered the redistribution of tone IV, it can be argued that this redistribution occurred in multiple waves. Pingyao, for example, retains some (but not all) of the closed syllables of middle Chinese, and exhibits a sandhi process that is unique to the reduplication of closed-syllable nouns (example (28), Pingyao sandhi rule 6), despite exhibiting an isomorphism in categories with Beijing in all other respects. Also, while there is at this point insufficient data to determine more specific details regarding the modern correlates of MC tone IV cognates in Pingyao vs. other languages, there are a fairly large number of anomalies in the distribution of MC tone IV words in Pingyao compared with others. For example, *ɕiq* in Pingyao ('to learn', 学), is in the modern TIII category, /53/, whereas we would expect it to correspond to *xue* in Beijing, which is in the modern TIb category, /35/, (modern TIb is /13b/ in Pingyao) if we assumed a single wave of tone IV redistribution. The two stem from the protoform **ɣak* in MC. In addition, the very fact that a syllable-initial voiced-obstruent/sonorant distinction conditioned a tone IV merge to categories III and Ib, respectively, whereas syllable-initial voicelessness conditioned a redistribution to all four tones, suggests two distinct processes at different points in time.

These changes can be broadly summarized in the table below (adapted from Chen 1976, 2000), which shows the modern Beijing Mandarin correlates for each of the MC tonal categories:

| MC onset | MC tones | | | |
|---|---|---|---|---|
| | I | II | III | IV |
| voiceless | 55 | 214 | 51 | 55, 35, 214, 51 |
| sonorant | 35 | 214 | 51 | 51 |
| voiced obstruent | 35 | 51 | 51 | 35 |

Table 14: Beijing Mandarin tonal correlates in MC.

It was mentioned above that there is evidence for the possibility that Proto-Tianjin experienced an upward drift of MC tone Ib and a downward drift of Ia. There are a few sources of evidence for this. The first is that the Tianjin correlate of MC tone II, /213/, still exhibits fairly strong similarity to the same tone in Beijing (/214/), Nantong (212), and Yantai (/214/). Although the Pingyao correlate of this tone exhibits a markedly different phonetic shape (/53/), its allotonic variant (/535/) possesses the same contour (concave), and it is reasonable to assume that the Pingyao version of MC tone II is the outlier of the three and experienced an upwards drift (/535/) before being simplified (/53/), leaving only an allotonic variant of the formerly concave tone.[54] If there was no significant movement of tone II, but the contrastive reanalysis caused a shift in the relationship between tone II and tone Ia (which are both low in modern Tianjin), there is

54 This does not imply that the Middle Chinese value for this particular tone was low and concave, but rather that it is likely that the tone category that was tone II in Middle Chinese was low and concave at a later point immediately prior to the four languages here diverging.

circumstantial evidence that tone Ia drifted downwards. This evidence is strengthened by another observation. Tone Ia and tone Ib in each language are the modern correlates of a single tone, (I) in Middle Chinese, which underwent a register split that was conditioned by syllable-initial voicing. Previous research by Mei (1970) suggests that Middle Chinese tone I likely had two allotonic variants, and the abundant literature on tonogenesis (e.g. Matisoff 1973) suggests that initial voicing generally conditions a low tone whereas initial voicelessness conditions a relatively higher tone. We thus expect that the MC tone I split likely resulted in a high and a low variant of the original tone I. In fact, the modern correlates of MC tone Ia (the voiceless variant) are higher or possess a higher allotone than the modern correlates of the intially voiced MC tone Ib ([55] vs [35] in Beijing; [13] and [31] vs. [13] in Pingyao; [31] and [35] vs. [31] in Yantai; [31] vs [24] in Nantong), in all languages **except** for Tianjin, in which the opposite is the case ([21] vs [45] for tones Ia and Ib respectively), suggesting a high degree of drift in Tianjin in these two categories.[55]

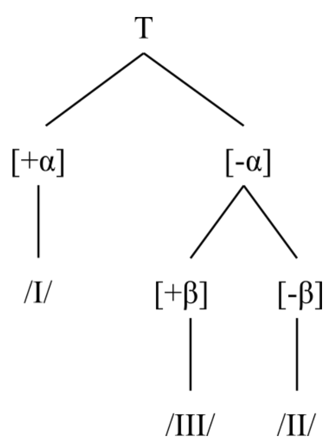
### 3.2.7.2 A Diachronic Pathway

The evidence thus far discussed provides a means for hypothesizing a specific pathway of the various diachronic changes of relevance to this study (this is mapped out in full in Appendix A). To start with, we know that the MC tone II partial merger with tone III (67b) likely occurred

55 The same general pattern can be observed in Yue languages, as in §4 below.

first. We see its effect homogenously in all five languages, and Chen (1976, 2000) states that the effect of this merger is not limited to Mandarin and Jin languages but extends well beyond the confines of the Mandarin group, something that is not universally the case for the other processes he outlines.

(68) Early Middle Chinese prior to full tone I split.



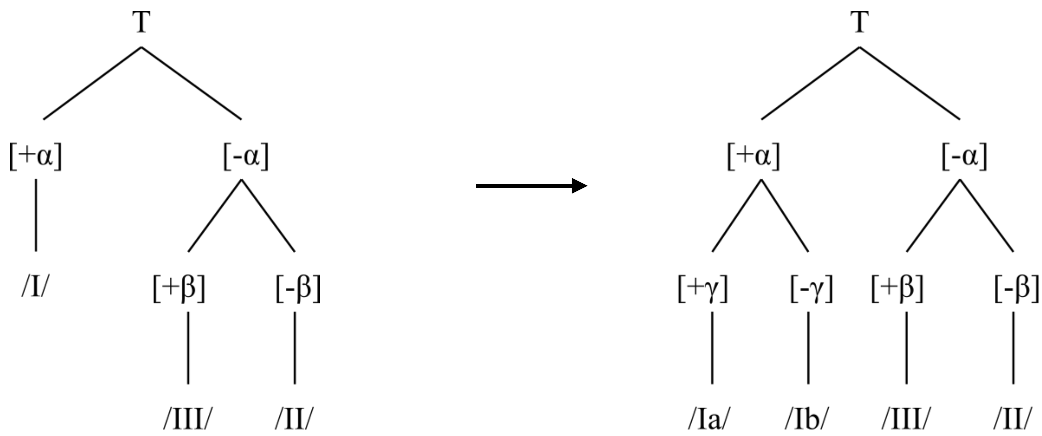

Minimality considerations (see §5 for a full discussion) allow us to very straightforwardly postulate an Early Middle Chinese contrastive hierarchy. Early Middle Chinese possessed three contrastive tones: TI, TII and TIII. A three-tone system has only one possible structural configuration (assuming no concrete feature labelling), [α] > [β], with one value of [α] resulting in a terminal node and both values of [β] resulting in terminal nodes. Tone I later bifurcated. Ko (2011) proposes that full category mergers only occur between phonemes which differ by a single feature at the most subordinate level. Oxford (2015) formalizes this as the *Sisterhood Merger Hypothesis*, which states that structural mergers apply to contrastive sisters. One should assume also that category splits, such as the MC tone I → Ia/Ib split, involve a split at terminal nodes, resulting in contrastive sisters. There are only two structural possibilities of a four-tone system where any two given tones are contrastive sisters: a symmetric, two-feature system, [α] > [β], with sisters possessing each value of [β] and sharing a value for [α], or a three-feature system [α] > [β] > [γ] with the contrastive sisters being specified for all three features. Given that all four-tone Mandarin languages are symmetric, two-feature systems, as discussed above, we postulate a symmetric, two-feature, four-tone Middle Chinese system immediately after the bifurcation of Tone I. Minimality considerations thus lead to the hypothesis that Early Middle Chinese Tone I was located at a terminal [α] node, and tones II and III were contrastive sisters, specified for [β], as above. Interestingly, it is good for the above analysis that this accords nicely with Chen's (2000) description of a partial TII → TIII merge (in (67b) above) which, again following Ko (2011), we expect between contrastive sisters. [56]

While the tone I split is not common to as wide a distribution of languages as the Tone II partial merge, it is common to both the Mandarin-Jin and Yue families (the latter of which is the subject of §4) whereas significant differences exist in the distribution of formerly tone IV syllables amongst the Mandarin and the Yue languages. As discussed above, there is evidence that the tone IV tripart redistribution likely occurred in waves, and while it is possible that the initial wave of redistribution co-occurred with the MC tone I split, we treat it as a single process which was fully realized after the full realization of the tone I split:

---

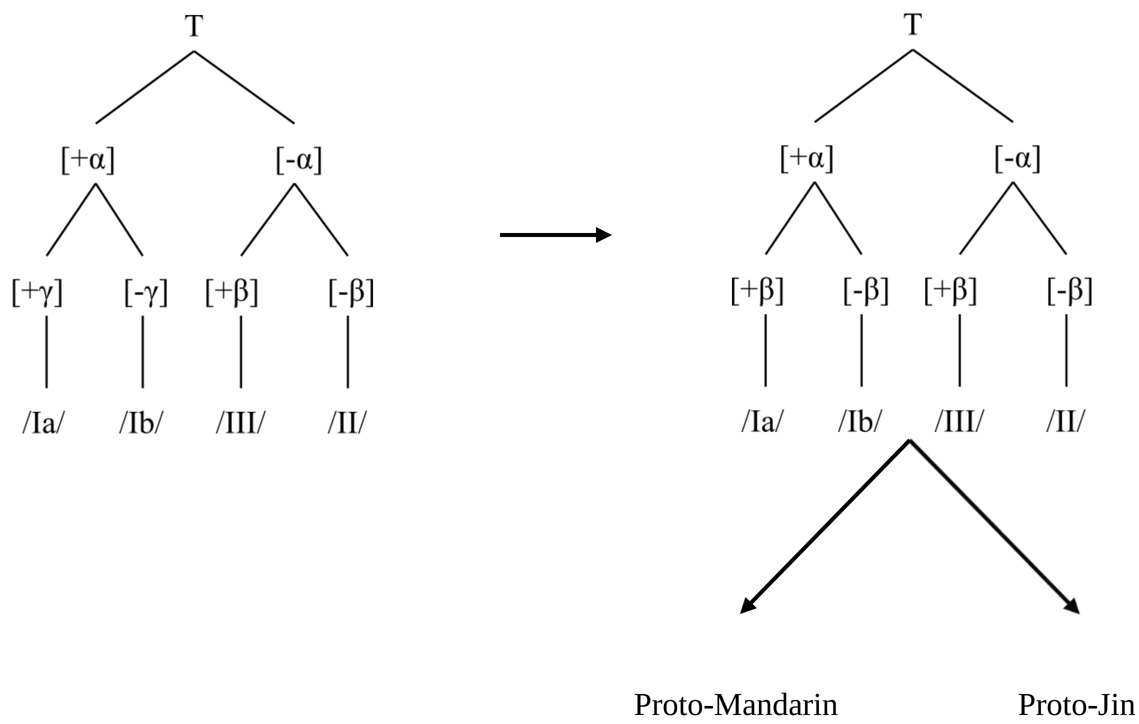
[56] Although Ko discusses full mergers and the loss of contrast, we would expect partial mergers to follow a similar logic. We do not rule out the possibility altogether of mergers between categories which differ only by a single feature but which are not contrastive sisters, but we would expect this to be comparatively rare. The partial merger in Yantai from Ib → III either occurred despite a difference in the superordinate [α] feature, rather than [β], or occurred alongside a reranking of features.

(69) Early Middle Chinese to Middle Chinese, tone I split




Because all Mandarin and Jin languages analyzed here are symmetric, four-tone, two-feature systems (with the exception of Nantong), we postulate that Proto-Guan reanalyzed the [γ] feature as a [β] feature (this is not the case in the Yue family of languages discussed below). Also, as discussed above, further evidence for the hypothesis that the tone IV redistribution occurred last of the three processes is found in the somewhat anomalous distribution of MC tone IV words in Pingyao vs. the other Mandarin languages investigated here. This discrepancy in the distribution of formerly tone IV words also provides further, albeit somewhat anecdotal, evidence for the bifurcation of the main Mandarin tree into distinct Mandarin and Jin subfamilies (Chen 2000; Li 1985; Wurm et al. 1987), which is otherwise a somewhat contentious division. Thus we postulate that Proto-Guan bifurcated into major Mandarin and Jin branches:

(70) Middle Chinese to Proto-Guan to Proto-Mandarin/Jin




After the Jin language divergence from the main Mandarin branch, Pingyao (Jin) likely retained more or less the same contrastive specifications until the present. Following the Mandarin branch however, Yantai likely branched off first, evidenced by the fact that, like Pingyao, it exhibits a citation tone merge in MC Ia and Ib, yet possesses the same tone IV cognate categories as Beijing and Tianjin. At some point after splitting, the Yantai Ib → III merge and contrastive reanalysis took place. This leaves Beijing and Tianjin in the main Mandarin branch, which all exhibit a full Ia / Ib split in citation form. Tianjin subsequently split off and underwent its own contrastive reanalyses discussed above. The precise placement of Nantong in this sequence is difficult to discern. We assume that it bifurcated prior to the Proto-Guan reanalysis of [γ] as [β], which would account for the structure of Proto-Nantong Stage 2 in (65), however, most analyses treat it as a Mandarin language proper, which would imply bifurcation post-reanalyses, and likely therefore after the Mandarin-Jin split. This would entail additional restructuring. The entire sequence of splits are depicted in Appendix A.

Although a detailed description of the phonetic shape of Middle Chinese tones remains elusive, we can make some attempts at determining some aspects. As discussed above, the two categories resulting from the tone I split, Ia and Ib, were likely high and a low respectively (although we have little information which would allow us to determine if they were contoured). Due to the fact that all the modern correlates of MC Tone II are either concave or possess a concave allotone (Pingyao), it is reasonable to assume that MC Tone II was also rising or concave, although we have insufficient information to determine its height value.[57] A larger database is necessary to make any more reasonable guesses.

## 3.3 Summary

In this chapter I motivated contrastive hierarchies for five Mandarin languages, based largely on activity patterns. The relative configurations of each contrastive hierarchy enabled us to postulate the contrastive specifications of Early Middle Chinese, from which I derived a diachronic pathway for each of the respective languages. Activity patterns also enabled us to discover that

[57] If we include the Yue languages into this determination, we can note that while there are flat modern correlates to MC TII (Binyang, Taishan, and likely Yangjiang), most are rising (HongKong, Guangzhou, Cenxi, Zhongshan) and none are unambiguously falling.

Tianjin, Nantong, and Yantai each underwent at least one contrastive reanalysis. We were able to motivate this via principles of parsimony, minimality, and, in the case of Tianjin and Nantong, phonetic naturalness. While contrastive hierarchies constrain phonological activity via the relations and oppositions between inventory members, hierarchies are able to re-organize themselves when they become too far removed from a phonetically natural representation of phonetic shape.

# 4

# The Yue languages

## 4.1 Overview

The Yue family, like the Guan or Mandarin family, is a major branch descended from Middle Chinese. It is spoken mainly in Southeast China, concentrated in Hong Kong / Guangdong and Gaungxi. Compared with the Mandarin family of languages, Yue languages tend to exhibit larger tonal inventories, simpler tonal contours (for example, concave and convex tones, while common in Mandarin languages, are nearly absent from Yue languages), and much less activity: tone sandhi is almost absent from the Yue family as a whole, making this group a particularly good complement to the analysis of the Mandarin languages above. The regional distribution of Yue languages analyzed in the present study is given below.


Figure 2: Geographical distribution of Yue languages.

Most Yue languages also exhibit a type of tonal alternation, known generally as *pinjam*, 'changed tone', which is generally analyzed as a type of morphological derivation by tone change. This process is variously realized in different languages, and involves, for example, nominalization of a verbal action, familiarization or diminutivization of an object, event or state, and various other complex semantic nuances (Jurafsky 1988; Yu 2007), and often occurs more commonly in familiar or casual speech than in more formal registers. This type of alternation involves a change in meaning, and is thus morpho-lexical in nature (Yu 2007). Kam (1977) analyzes a subset of *pinjam* as a type of derivation by tone change, arguing that the modern Cantonese system is an expansion of a similar system in classical Chinese. *Pinjam* are excluded from the present study as they are generally analyzed as involving the full replacement of tones without any purely phonological constraint on their distribution (see Bauer and Benedict 1997 for a detailed discussion).[58]

The following sections deal with Yue languages. Section 4.2 deals with the standard, prestige Yue language, Hong Kong Cantonese. Section 4.3 describes and analyzes a number of other Yue languages: Cenxi in §4.3.1, Zhongshan in §4.3.2, Binyang in §4.3.3, Taishan in §4.3.4, and Yangjiang in §4.3.5. Section 4.4 is a summary.

## 4.2 Hong Kong and Guangzhou Cantonese

There are two main dialects of so called ‘Standard’ Cantonese, the prestige Yue dialect: Guangzhou and Hong Kong, spoken in the port cities of Guangzhou and Hong Kong, respectively. The tonal inventory of Guangzhou Cantonese is given in (71) below. The two varieties are mutually intelligible and both phonetically and phonologically very similar, though easily distinguished tonally by the presence or absence of the high falling tone, (i), which is present in Guangzhou Cantonese, but has completely merged with the high flat tone, (ii), in Hong Kong Cantonese.

---

[58] This being said, the fact that /55/ in Hong Kong does not participate in *pinjam* hints at the possibility that it may be more complex than commonly treated in the literature. I leave this to future study.

(71) Hong Kong/Guangzhou Cantonese[59] (adapted from Barrie 2007, Bauer and Benedict 1997)

| | | | | | |
|---|---|---|---|---|---|
| i. | *saan* | high falling | / 53 / | 'to close' | (Guangzhou only) |
| ii. | *gʌm* | high flat | / 55 / | 'gold' | |
| iii. | *hou* | low, high-rising | / 25 / | 'good' | |
| iv. | *sei* | mid flat | / 33 / | 'four' | |
| v. | *yau* | low falling | / 21 / | 'oil' | |
| vi. | *ngo* | low, mid-rising | / 23 / | 'I, me' | |
| vii. | *daai* | low flat | / 22 / | 'big' | |
| viii. | *sʌp* | high checked | / 5q / | 'wet' | |
| ix. | *baat* | mid checked | / 3q / | 'eight' | |
| x. | *dip* | low checked | / 2q / | 'dish' | |

A pitch trace of the six full citation tones for Hong Kong Cantonese is given below:

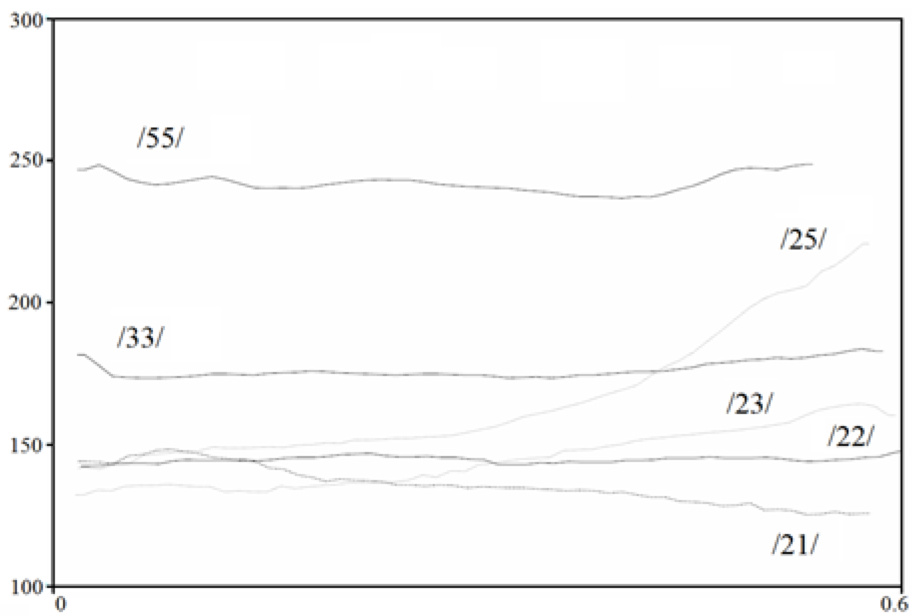



Figure 3: Hong Kong Cantonese citation tones (adapted from Mok and Wong 2010)

59 Bauer and Benedict (1997) differ in their numbering as follows: i. /52/, iii. /35/, viii. /55/, ix. /33/, x. /22/.

### 4.2.1 **Checked Tones**

It is general consensus that, although the checked tones (viii-x) are, by convention, treated as separate full categories in traditional Chinese dialectology (Chen 2000:16), they may be analyzed as being allotonic with corresponding full tone counterparts (Bao 1999; Barrie 2007; Bauer and Benedict 1997; Chen 2000; Yue-Hashimoto 1972). Checked tones occur only on obstruent-final syllables, and /5q/ and /3q/ are argued to be in complementary distribution with tone /33/, and tone /2q/ with tone /22/ (Chen 2000:16). /5q/ and /3q/ have been argued to be distinguished by a tense-lax distinction (Yue-Hashimoto 1972, Chen 2000) with the former occurring on short, lax vowels, and the latter on tense, long vowels, as in Table 15 below (adapted from Chen 2000).

| MC tone | Modern tone | | | |
|---|---|---|---|---|
| IIIa | 33 | *si:* 'to try' | *fu:*, 'rich' | *wʌn*, 'to shut up' |
| IIIb | 22 | *si:* 'affairs' | *fu:*, 'father' | *wʌn*, 'to transport' |
| IVa | 5q | *sɪk*, 'to know' | *fʌt*, 'sudden' | *wʌt*, 'twisted' |
| IVa | 3q | *sɛːk*, 'lead' | *fa*ː*t*, 'law' | *wa*ː*t*, 'to dig' |
| IVb | 2q | *sɪk*, 'to eat' | *fʌt*, 'to punish' | *wʌt*, 'pit (fruit)' |

Table 15: Three sets of Cantonese morphemes, by tone.

It should be noted that although this tense/lax distinction may appear to be phonologically active in Cantonese, Barrie (2003) points out that vowel quality for most vowels is predicted by complex co-occurrence relations, with lax vowels (/ɪ, ɛ:, ʊ, œ:/) occurring before /k, ŋ/ and tense vowels (/i:, e, u:, o, ø/) occurring before /w, m, n, p, t/. Exceptions to this tendency are /y:/, which has no allophonic variant but only occurs in open syllables and before /n/ and /t/; lax /ɔ:/, which occurs before both velar and non-velar consonants, and the low vowels (/ʌ/ and /a:/), which are contrastive for length and can both co-occur with any coda. Complicating the picture is the fact that, while each of the five vowels which comprise a tense-lax allophonic pair has one member long and the other short, length is not predictable from tenseness. The vowel system in Cantonese is the subject of significant disagreement, however, if Chen's (2000) data (from Yue-Hashimoto 1972 and Yuan 1960) is correct, /3q/ and /5q/ are likely predictable by vowel length, rather than tenseness. Barrie (2003) comments that length, aside from in the low vowels and in /y:/ cannot be predicted in any obvious manner.

### 4.2.2 High Tone Merge and Sandhi

The merge of the two high tones in Hong Kong relative to Guangzhou is exemplified in the following example of Guangzhou minimal pairs versus Hong Kong homophones (adapted from Barrie 2007):

(72) *saːn*, /55/ 'hill' *saːn*, /53/ 'to close' (Guangzhou)
*saːn*, /55/ 'hill' *saːn*, /55/ 'to close' (Hong Kong)

Bauer and Benedict (1997:120) state that while "most Guangzhou speakers make a distinction between High Falling (i.e. /53/) and High Level (i.e. /55/) tones; most Hong Kong speakers, on the other hand, have lost the High Falling tone (or do not contrast High Falling with High Level) and instead use the High Level tone. In other words where Hong Kong speakers have High Level, Guangzhou speakers have either High Falling or High Level" (1997:120).

A precise account of the development of the high tones in Cantonese does not exist in the literature. Whitaker's 1955 paper on a class of morphologically conditioned tone changes argues that Standard (Hong Kong) Cantonese possessed ten tones (i.e. the 7 citation tones and 3 checked tones of Guangzhou), whereas McCoy's 1966 descriptive work on some 24 distinct subdialects of Cantonese puts the number of citation tones of Hong Kong Cantonese at 6, with the additional three checked tones. Interestingly, McCoy lists Hong Kong Cantonese as possessing a high falling tone, /53/, with no high flat tone, however he does list /53/ as possessing an allotonic flat variant in closed syllables (likely modern Cantonese's /5q/). Similarly, an earlier (1912) work by Daniel Jones and Kwing Tong Woo lists six main tones, among them an 'upper falling' tone which possesses an 'uppermost level' variant. Further confounding this issue is the argument that some speakers possess two high tone variants (Yue-Hashimoto 1973:178-180), with the high falling variant as the underlying form. Yue-Hashimoto argues that, depending on speaker and context, some speakers may utilize, for example the high flat tone for almost all nouns, some for a lesser number of them, and some speakers use the two tones interchangeably for the same morpheme. She goes on to describe the existence of a large number of speakers who utilize the high flat tone for nouns and categories other than nouns which are familiar in nature or have some familiar relationship to the speaker, and the high falling tone for other morphemes, although the reverse is not true. In any case, it seems strongly likely that there have been multiple concurrent variants for some time, and indeed McCoy argues that some of the discrepancies found in older descriptions of 'Standard Cantonese' can be accounted for by one variant or another being chosen to represent standard forms. The variation seen in descriptive accounts regarding the high tone(s) from at least the early 20th century is compatible with a general sound change culminating in a single category (high-flat) for modern Hong Kong Cantonese.[60]

Whereas the two high tones in Modern Hong Kong Cantonese are completely merged, Guangzhou Cantonese possesses a sandhi rule relating them, implying that the two high tones differ minimally, by a single feature.

(73) Guangzhou high tone sandhi (adapted from Bauer and Benedict 1997)

53 → 55 / __ {/53/, /55/}

| | | | | | |
|---|---|---|---|---|---|
| | *paw* | *ji* | | *paw* | *ji* |
| | | | → | ‘guaranteed cure’ | |
| | /53/ | /53/ | | **55** | 53 |
| *i.* | *tsan* | *tsy* | | *tsan* | *tsy* |
| | | | → | ‘pearl’ | |
| | /53/ | /53/ | | **55** | 53 |
| *ii.* | *ji* | *saŋ* | | *ji* | *saŋ* |
| | | | → | ‘doctor’ | |
| | /53/ | /55/ | | **55** | 55 |

60 Variation certainly still exists in this area in Hong Kong Cantonese. More broadly, while we utilize a single inventory of tonal digits to represent each language here, multiple realizations likely exist in many cases. We would expect inventory members specified with fewer features to, potentially, possess a broader possible range of variation (a similar argument is made in Rice 1995).

| | | | | |
|---|---|---|---|---|
| *iii.* | *kwan* | *taŋ* | | *kwan* *taŋ* |
| | | | → | ‘turn off a light’ |
| | /53/ | /55/ | | **55** 55 |

### 4.2.3 Cognate categories and feature representation

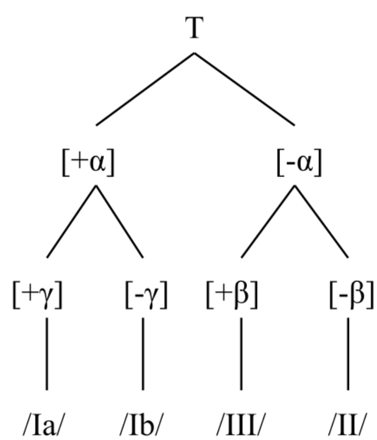
As will be seen below (Table 16), these two tones also derive from the same MC tone, providing further evidence that these two tones differ minimally. In fact, the modern Cantonese tones are all derived from their Middle Chinese counterparts in a phonetically transparent manner. This suggests either a remarkable lack of tonal drift over time, relative to much of the Mandarin inventories investigated in this study, whose modern tones have at best a tenuous phonetic relationship with their MC categories, or a synchronization of tonal drift of all categories based on contrastive features, where the drift of a single tone pushes or pulls other tones with it to preserve contrastive structure and inventory relations.

| MC Onset | MC TI | MC TII | MC TIII | MC TIV |
|---|---|---|---|---|
| voiceless | /53, 55/ (Ia) | /25/ | /33/ | /5q, 3q/ |
| voiced | /21/ (Ib) | /23/ | /22/ | /2q/ |

Table 16: Cantonese Tones by Middle Chinese categories (Bauer and Benedict 1997, adapted)

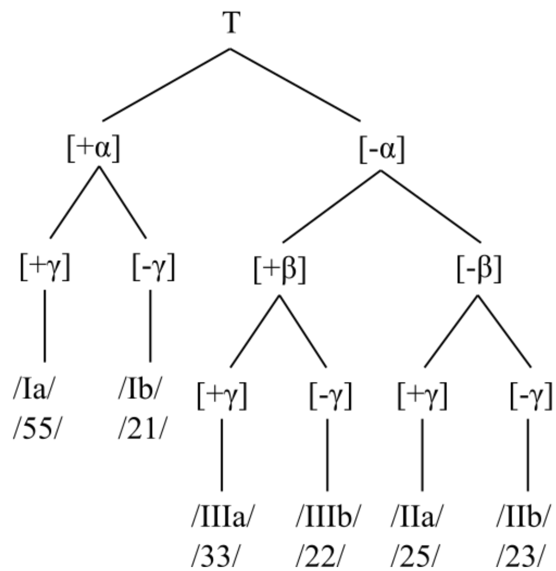
As discussed previously, MC tones II and III were contrastive sisters at the time of the redistribution of tone IV and the bifurcation of tone I into contrastive Ia and Ib categories. This bifurcation occurred via a register split into high and low categories due to the loss of a voice contrast in syllable-initial position, with syllables with formerly voiced initials being reanalyzed as possessing a contrastively [-high] tone (Ib), and those with voiceless initials reanalyzed as [+high] (Ia).

(74) Late Middle Chinese contrastive hierarchy




While the Tone I split is shared with most languages descended from Middle Chinese, in most Yue languages tones II and III also bifurcated on the basis of height (i.e. into IIa, IIb, IIIa and IIIb, respectively). Comparing Table 16 and the hierarchy in (74), based on phonetic shape and on what we know about the likely structure and development of Middle Chinese, we might postulate a fairly straightforward outline of a feature tree for Cantonese as an expansion of the Late Middle Chinese contrastive hierarchy above.

(75) Hong Kong Cantonese, [61] [α] > [β] > [γ]




With this structure, [α] separates the modern correlates of Tone I from the rest of the tree, [β] separates the Tone II correlates from the Tone III ones, and [γ] represents the Yue register split, dividing the remaining categories into high and low. Regarding the structure of Middle Chinese in (74), we might also postulate a [β] feature rather than a [γ] one in the domain of [+α] since we know that the tone I bifurcation, shared with the Mandarin/Jin languages, resulted in a four tone symmetric system. Although we might assume that, for purposes of information economy, the feature dividing MC Ia and Ib is the same as the one dividing II and III, it is perhaps more conservative to postulate distinct features in the absence of evidence to the contrary (unlike, for

---

61 There are reasons to suspect that the hierarchy in (58) was subject to a subsequent structural reanalysis, and should thus be considered a late Proto-Hong Kong Cantonese. The evidence for this is beyond the scope of this section and is discussed in more depth in §5.4.2.

example, Beijing Mandarin, which possesses evidence from activity). We know that tone I bifurcated on the basis of height due to devoicing, and it would be unlikely for this separately motivated process to correspond to the extant feature dividing tones II and III. As well, we do know that tones I, II, and III all bifurcated on the basis of height via the same process, and although this likely happened over a fairly long period of time (evidenced by the fact that only tone I split in the Mandarin family whereas all three split in Yue), postulating a single feature, [γ], to account for this captures this fact.

Although (75) represents an intuitive structure that captures both the phonetic shape of modern tones and their historic categories, we might initially be somewhat suspicious of such an analysis, knowing that the Mandarin languages discussed in Chapter 3 do a fairly poor job of settling themselves into phonetically sensible structures, and Cantonese has no evidence from phonological activity to support the structure in (75). However, some additional facts about Cantonese lend support to just such a phonetically transparent configuration.

A number of studies have reported partial mergers between tones in some speakers or populations which are in addition to the merger of the two high tones in HK Cantonese discussed above. Kei et al. (2002) report on a merger between /25/ and /23/ amongst six out of fifteen participants in a study on perceptual judgement and instrumental measure of production.[62] Similarly, So (1996) reports on a study by So and Varley (1991) finding that experimental subjects

[62] Kei et al. (2002) call them errors of perception or production, but the purpose of their study, as Bauer et al. (2003) point out, was to provide a normative profile of 'Standard Cantonese' tones for use by speech therapists, rather than to describe variants in a community.

also often confused the two rising tones. Bauer et al. (2003) report partial mergers between the same two rising tones, /25/ and /23/, in some Hong Kong speakers, arguing it is likely sociolinguistically-based variation. They argue that some Hong Kong Cantonese speakers do not distinguish between the two tones, but tend to favour one or the other or even one intermediate between the two, and they postulate multiple tonal subsystems within the Hong Kong speech community. A production study by Mok and Wong (2010) investigated multiple potential merger contexts, again finding evidence for a merger between /25/ and /23/, and interestingly enough, also for a merger between /33/ and /22/. The balance of evidence thus supports postulating that in Hong Kong Cantonese, /25/ and /23/ are contrastive sisters, differing by a single feature at the most subordinate level, and /33/ and /22/ are also contrastive sisters. This is precisely what one would expect given the historical origins of these tones, but the transparency of the phonetic relationship between modern tones, their features and structure, and their historic origins is remarkable in this language. We can thus argue for the straightforward contrastive hierarchy for Hong Kong Cantonese postulated in (75) above.

Our [α] feature is best analyzed as [peripheral], with those tones with positive values having pitch values at the peripheral upper and lower pitch ranges.[63] [β] is analyzable as a flat/rising distinction. Our [γ] feature is thus a simple height distinction, with positive values being [high] and negative values being [low] or [-high].

---

63 Therefore, /22/ is viewed as simply low, whereas /21/ is extra low and contrastively flat.

Since modern Hong Kong Cantonese possesses only a high flat tone, and older accounts describe either only a falling high tone, or both a falling high and a flat high tone, but no older descriptions describe **only** a flat high tone, it is likely that the existence of only the high flat tone in Hong Kong Cantonese is a fairly recent development resulting from a merger of the two high tones, and varieties which possessed both a falling and a flat tone resembled modern Guangzhou Cantonese, which still possesses both. We can postulate the following sequence of changes to account for all attested varieties.

(76) Tonal changes from Late Middle Chinese to modern Cantonese

a) ***Register Split (for Tones I, II, III)***

Tone → { a (high) / voiceless initial<br>b (low) / voiced initial

b) ***Tone Ia split***

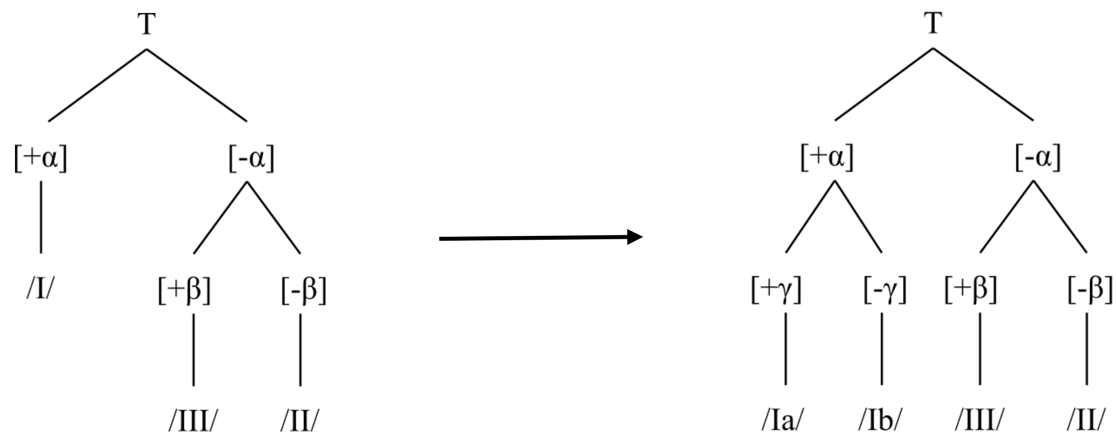
Tone Ia (high) → { Tone Ia.i (flat)<br>Tone Ia.ii (falling)

c) ***Tone Ia merge***

Tone Ia.ii → Tone Ia.i (flat)

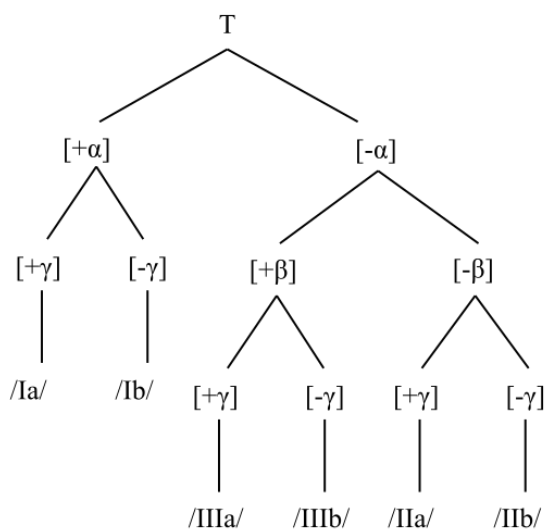
Applying these to our Middle Chinese contrastive hierarchy yields the following sequence of trees:

(77) Early Middle Chinese Late Middle Chinese / Proto Yue



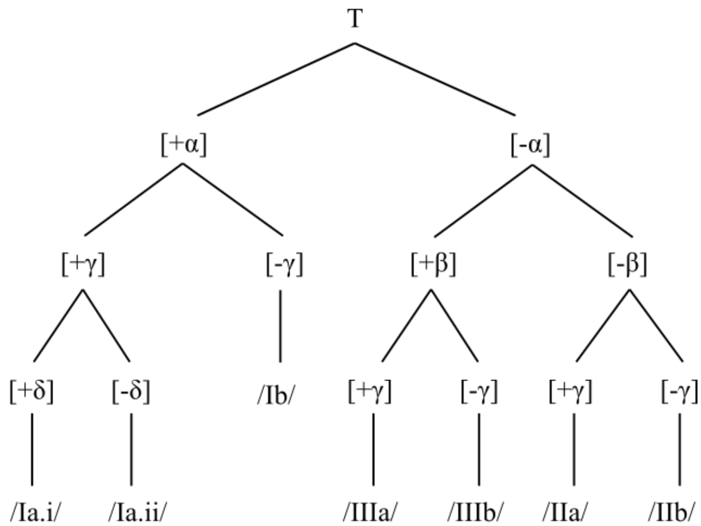

As discussed above, we know that the TI (shared with Mandarin/Jin languages) split preceded splits of TII and TIII (specific to the Yue family). Applying the register split to all categories gives the following:

(78) Middle Cantonese (TII, TIII split)



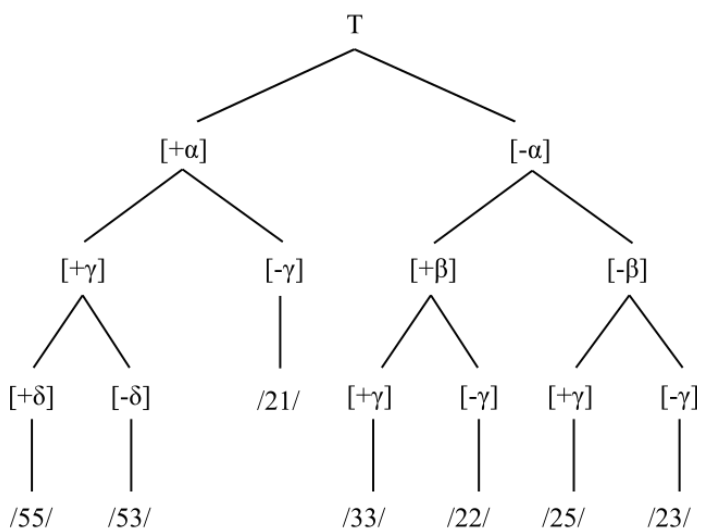

The tree in (78) represents a common ancestor to most Yue languages. The direct ancestor to both Hong Kong and Guangzhou is the result of the subsequent TIa split (i.e. (76b) above) applied to (78), we postulate, as below.

(79) Proto-Hong Kong/Guangzhou




The tree in (79) above is structurally identical to modern Guangzhou Cantonese:

(80) Guangzhou Cantonese, [α] > [β] > [γ] > [δ]




Further evidence supporting our Tone Ia split comes from cognate data: the class of words occurring with either a high flat or high falling tone in Guangzhou Cantonese is completely subsumed by the Tone Ia category (as in example (72) above). Modern Hong Kong Cantonese, (75), is therefore the result of the merger (76c) between TIa.i and TIa.ii in (79), with the high flat tone as a target. Hong Kong Cantonese thus experienced a Tone I split (as did all modern Chinese dialects), a Tone Ia split, and a very recent Tone Ia merge, whereas Guangzhou Cantonese experienced only the Tone I and Tone Ia splits, and attested variants possessing only a single high falling tone are thus the products of a variant of rule (76c) with the falling, rather than flat, tone as a target, or are languages which did not experience the initial tone Ia split, as below.[64]

An implicit assumption can thus now be made explicit. Just as we assume, given the *Sisterhood Merger Hypothesis* (Oxford 2015, see also Ko 2011) discussed in §3.2.7, that structural mergers only occur between contrastive sisters, we can also assume the converse, namely that structural splits result in contrastive sisters. The modern categories resulting from the register split of MC tones I, II, and III result in contrastive sisters, and the contrast resulting from the Tone Ia split also manifests at the most subordinate level of the tree. Feature labelling of Guangzhou is identical to Hong Kong Cantonese, except of course for [δ], which is best analyzed as either [falling], or as [contour] in the contrastively high register (i.e. [+γ]), in line with the argument (Barrie 2007; Yip 2001), below, that a pitch contrast serves to determine contour directionality.

Barrie (2007), working from a similar framework as the present study, analyzes Guangzhou and Hong Kong Cantonese as having the following feature specifications (adapted):

---

64 One could possibly make an argument that /55/ in modern HK Cantonese is the simple result of drift from /53/ to /55/, but this requires postulating both drift and reanalysis ([+contour] → [-contour]), and does not easily fit into a larger picture involving systems with either of the two high tones or individual systems possessing both high tones (e.g. Guangzhou), and outright contradicts older accounts, discussed above, attesting two high tones in HK Cantonese specifically, and is thus a much less parsimonious analysis.

(81) [upper] > [high] > [contour]

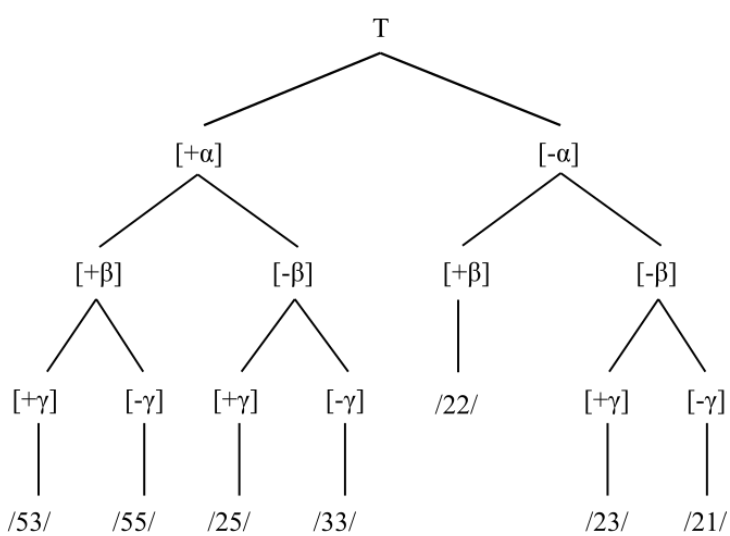
/55/ [+upper, +high] /33/ [+upper, -high] /25/ [+upper, -high, +contour]
/21/ [-upper, -high] /22/ [-upper, +high] /23/ [-upper, -high, +contour]
/53/ [+upper, +high, +contour] (Guanghzou Cantonese only)

Barrie makes the argument, echoing Yip (2001), that directionality in [contour] features may be implicitly specified via a register contrast, with [+high, +contour] tones being falling, and [-high, +contour] tones being rising. Barrie also argues that /21/, rather than being falling, is an underlying ‘extra low level’ tone, and is unspecified with regards to [contour]. Using abstract features with Barrie’s labelling to serve as a basis for comparison, we obtain the following:

(82) Cantonese feature specification in Barrie (2007), abstract features, [α] > [β] > [γ]

/55/ [+α +β] ([-γ]) /33/ [+α -β] ([-γ]) /25/ [+α -β +γ]
/21/ [-α -β] ([-γ]) /22/ [-α +β] /23/ [-α -β +γ]
/53/ [+α +β +γ] (Guangzhou Cantonese only)

(83) Cantonese feature specification in Barrie (2007), tree form.




We might however disfavour this analysis for two reasons. Firstly, the partial merger reported in Mok and Wong (2010) between /33/ and /22/ would be exceedingly unlikely if the pair shared no features. As with sandhi, we would by default expect mergers to occur between pairs which differ by a single feature, or better, between contrastive sisters. Secondly, assuming

(81, it becomes exceedingly difficult to account for the diachronic development of the modern Cantonese tonal system: /55/ and /21/, as well as /33/ and /22/, which stem from single categories, share no features with each other. Given that we also know that the TI, TII, TIII split into high and low variants and the TIa split into TIa.i and TIa.ii occurred most recent1y relative to the genesis of other categories we would assume that, barring reanalyses, these categories would exhibit structural relations with each other at the most subordinate level. That is, IIa (/25/) and IIb (/23/), IIIa (/33/) and IIIb (/22/), and TIa.i (/55/) and TIa.ii (/53/) should be contrastive sisters, and TIa.i (/55/), TIa.ii (/53/) and TIb (/21/) should be similarly related, again barring reanalyses. This is not the case in Barrie's contrastive specifications, and only the two high tones exhibit the type of structural relations we would expect given the evidence. A tree like that in (83) might be possible in a system that has undergone massive tone shift and radical restructuring, but there is no evidence for this in Cantonese. Feature encoding and potential reanalyses are revisited in §5.4.2.

Compared with the Mandarin languages, Yue languages have a relative paucity of tone sandhi (a single process in Guangzhou and none at all in Hong Kong Cantonese), a generally larger tonal inventory (with some exceptions, discussed below), and relatively simpler tonal shapes. Hong Kong and Guangzhou are also more amenable to a transparent phonetic labelling which also preserves their respective diachronic development from Middle Chinese.

The reason for this may lie in the task that learners face when acquiring the language. Consider the two possible extremes of a system with absolutely no phonological processes vs. one with more processes than inventory members. In a system with no phonological processes whatsoever, the only criteria by which learners may form hypotheses about a phonology is the phonetic content of the inventory itself. Phonetic drift in such a system can affect the learnability of the system in two possible ways: towards a more easily learned system (i.e. a more phonetically natural mapping of features to phonetic content, evenly distributed inventory members etc.), or towards a less easily learned system.[65] This eventually leads to islands of stability from which any drift yields less easily learned systems, which are thus less likely to be acquired, constraining further drift. Drift in such a scenario can obviously still occur, but we would expect it to be motivated by factors external to our particular tonal subsystem, which would need to exert enough of an influence on our subsystem to overcome this relatively high degree of stability. In contrast, in a system with a large number of phonological processes, learners form hypotheses about phonological structure based on these processes. If two inventory members are linked via a sandhi rule, the link itself provides a direct and explicit structural association between the categories, without any requisite regard to phonetic content of the categories themselves. This may liberate the category in question to drift more freely as phonetic content does not serve as the only determiner of inventory relations, and the drift of a single tone does not threaten to either drag other tones along with it to preserve structural relations or radically restructure our entire

[65] While ‘drift’ is utilized here to refer to any gradual phonetic change of a single inventory member, there are a number of distinct senses of the term: i) motion away from a phonetically natural arrangement motivated by external factors; ii) compensatory motion towards a phonetically natural arrangement motivated by system-internal learnability considerations in response to external factors (similar to ‘push’ and ‘pull’ in chain shifts in the sense of Martinet (1952)) ; iii) incidental motion towards a phonetically natural arrangement motivated by external factors; iv) the gradual reinterpretation of underlying representations from natural variation in the distributions of inventory members to more learnable systems, as above. ii) and iv) may not be wholly distinct in all cases.

inventory.[66] We would therefore expect only factors such as a relatively high degree of perceptual similarity between an inventory member and a given region into which another tone may potentially drift, for example, to constrain the drift that may occur in these systems. This is revisited in §5.

## 4.3 Other Yue languages

The following sections focus on motivating contrastive hierarchies for several other Yue languages. Some, like Standard Hong Kong Cantonese, descend from Middle Cantonese (Cenxi, Standard Taishan, Doushan Taishan), whereas others bifurcated earlier, and are descended from Proto-Yue (Binyang, Yangjiang, Zhongshan). These relationships are laid out in detail in Appendix A. None of them exhibit any tone sandhi, and all are characterized by having fairly transparent phonetic substance with respect to the diachronic origins of their respective categories.

### 4.3.1 Cenxi

Cenxi is a small, relatively isolated city in a mountainous region located approximately 150 km west of Guangzhou. Lee (1993) writes that it is one of the oldest Chinese settlements in

---

66 This entails that tone sandhi is acquired in concert with individual tones. There is evidence that this is the case, at least in Mandarin, the only Chinese language with robust sandhi for which detailed acquisitional data exists. Li and Thompson's (1977) systematic study on the acquisition of tone in Mandarin found that the tonal system is generally acquired relatively early, and tone sandhi is rapidly and fully acquired as soon as the child begins to make two word utterances, well before the disappearance of replacement errors. Hsu (1987) reports similar findings regarding order of acquisition. Singh and Fu (2016) hypothesize that tone sandhi directly influences acquisition of the tones it is associated with in Mandarin, arguing that while categories emerge early, they take much longer to 'crystallize' and reach adult-like forms.

the Beilu River valley. Like Hong Kong and Guangzhou Cantonese, Cenxi possesses a relatively conservative tonal system with phonetically transparent categories. Like HK Cantonese, it is a six-tone system:

(84) Cenxi (adapted from Tsuji 1977)

| | | | | |
|---|---|---|---|---|
| i. | *kei* | high falling | / 52 / | 'chicken' |
| ii. | *po* | low falling | / 21 / | 'wife' |
| iii. | *tu* | mid-rising | / 35 / | 'vinegar' |
| iv. | *tou* | low-rising | / 12 / | 'abdomen' |
| v. | *po* | high flat | / 44 / | 'break' |
| vi. | *ʃei* | low flat | / 22 / | 'pledge' |

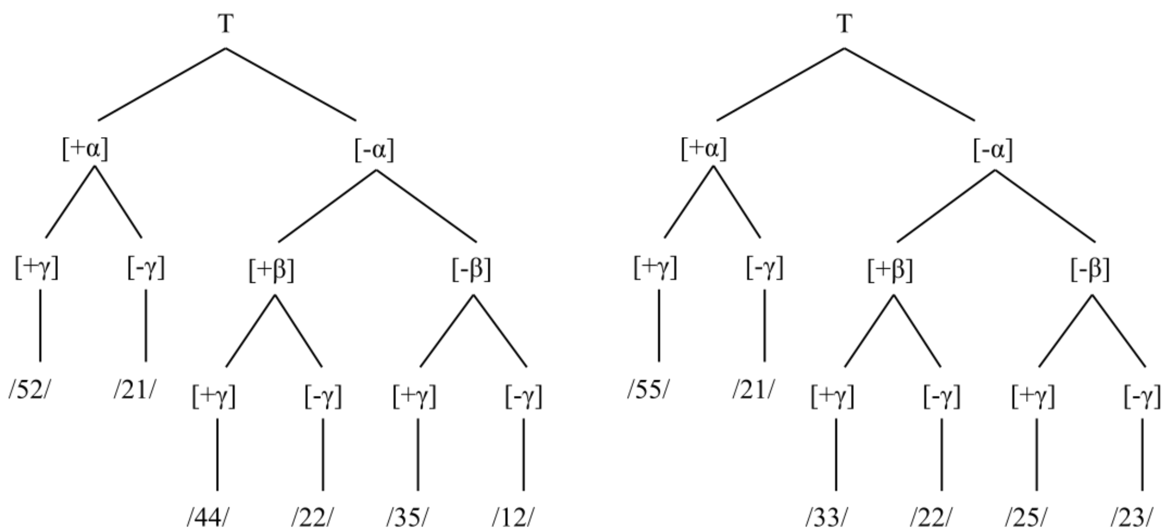
As with other Yue languages, Cenxi also possesses a set of checked tones allotonic with its respective full-tone counterparts. One notable difference is that /35q/ is in complementary distribution with the modern correlate of MC TIIa, rather than MC TIIIa. The correspondence between modern tones and the MC categories is given below:

| MC Onset | MC TI | MC TII | MC TIII | MC TIV |
|---|---|---|---|---|
| voiceless | /52/ (Ia) | /35/ | /44/ | /5q, 35q/ |
| voiced | /21/ (Ib) | /12/ | /22/ | /2q/ |

Table 17: Cenxi (adapted from Chan 1980)

Given what we know about HK and Guangzhou Cantonese, we can uncontroversially postulate a fairly straightforward tree for Cenxi as follows.

(85) Cenxi, [α] > [β] > [γ] Hong Kong Cantonese




As can be seen, Cenxi is more or less identical to HK Cantonese, with only minor differences in the height of various citation tones, except for the presence of a high falling, rather than a high flat [+α +γ] tone in Cenxi. For Cenxi, we can assume either the same set of diachronic processes for Hong Kong with a different merger target for our Tone Ia merge (falling rather than flat), or we can assume that Cenxi did not undergo a TIa split at all, and it is structurally identical to Middle Cantonese in (78). The latter is preferred here given the fact that Cenxi is even more phonetically transparent than HK/Guangzhou, suggesting a stable configuration in the absence of external impinging factors on the tonal subsystem. The most obvious feature labelling would be a falling/not falling distinction for [α], a flat/rising distinction for [β], and a high/low distinction for [γ].

### 4.3.2 Zhongshan

Zhongshan is a prefecture-level city immediately south of Guangzhou. The Zhongshan language has the fewest number of citation tones of all the Yue languages.

(86) Zhongshan (adapted from Chan 1980)

| | | | | |
|---|---|---|---|---|
| i. | *pan* | high flat | / 55 / | ‘to run away’ |
| ii. | *mun* | high falling | / 51 / | ‘a door’ |
| iii. | *pun* | low-rising | / 13 / | ‘root’ |
| iv. | *pan* | low-flat | / 22 / | ‘stupid’ |
| v. | *sɪk* | high-checked | /5q/ | ‘colour’ |
| vi. | *kit* | low-checked | /2q/ | ‘clean’ |

The correspondence between Zhongshan tones and their MC categories is given below:

| MC Onset | MC TI | MC TII | MC TIII | MC TIV |
|---|---|---|---|---|
| voiceless | /55/ (Ia) | **/13/** | **/22/** | /5q, 2q/ |
| voiced | /51/ (Ib) | **/13/** | **/22/** | /2q/ |

Table 18: Zhongshan (adapted from Chan 1980)

A number of properties are of interest. Firstly, although MC TIV underwent a redistribution in Zhongshan with the target tones determined by syllable-initial voicing, there is a subclass of formerly TIV words that had voiceless initials that were realized as low, rather than high. This has been taken as evidence that MCTII and MCTIII in Zhongshan previously split and then merged again (Lee 1993), causing anomalies in the distribution of modern checked tones with respect to their expected distribution based on syllable-initial voicing. Closer examination reveals however that the distribution of MC tone IV on formerly voiceless syllables is neatly predicted by Middle Chinese rhyme type in a reconstruction in Pulleyblank (1977) referenced in Chan (1980). Although it has been argued that Zhongshan, like HK Cantonese, presently has no active tense/lax distinction (Chan 1980:40), vowel length and quality in modern HK Cantonese do broadly predict which of /5q/ or /2q/ a tone in Zhongshan falls into.[67] We thus simply postulate an active tense/lax distinction at some earlier stage of the language shared by both HK Cantonese and Zhongshan, with the subsequent loss of this distinction conditioning a category change in checked tones occurring on formerly voiceless initials.

| Zhongshan tone | Zhongshan | HK Cantonese | |
|---|---|---|---|
| *5q* | *sik* | *sɪk* | 'colour' |
| *2q* | *kit* | *ki:t* | 'clean' |
| *2q* | *tit* | *ti:t* | 'decorum' |
| 5q | *fat* | *fʌt* | 'suddenly' |
| 2q | *wat* | *wa:t* | 'slippery' |

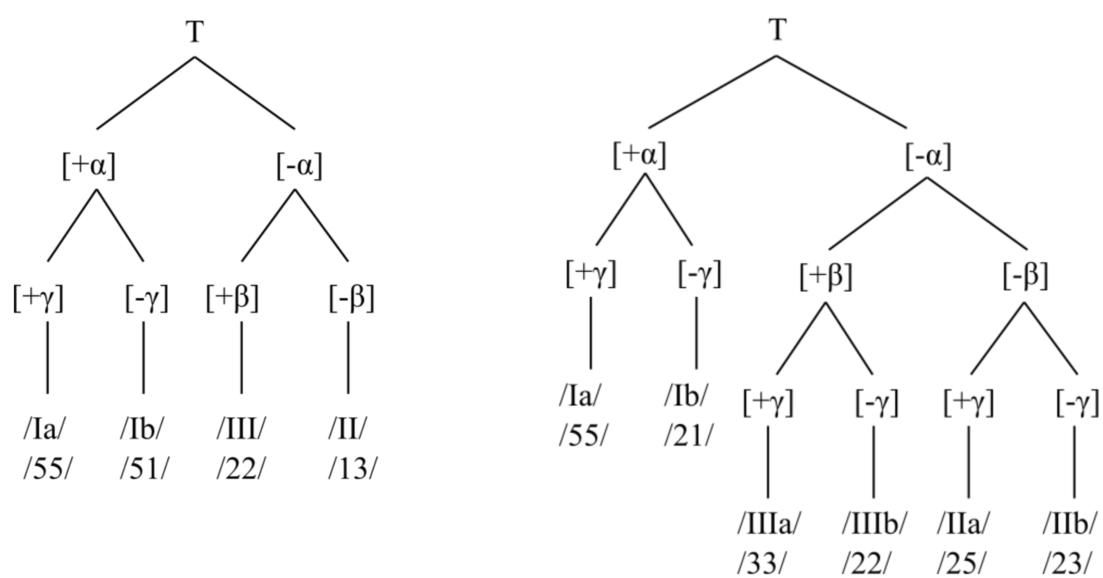
Table 19: Zhongshan checked tones by tense/lax vowel quality in Hong Kong Cantonese

A distinctive property of Zhongshan is the small tonal inventory. Unlike all other Yue dialects, Zhongshan does not exhibit a register split in either of the MC TII or TIII categories, making the tonal inventory of the language resemble the Mandarin family. A feature tree of

67 This is somewhat of a simplification, though Middle Chinese rhymes can be broadly derived from modern Cantonese forms. See Chen and Newman (1984a, 1984b, 1985) and Newman and Raman (1999) for more detailed discussion.

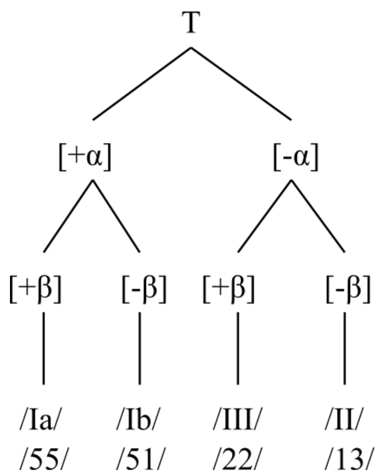
Zhongshan is given below. Interestingly, the contours of the modern correlates of TII and TIII in Zhongshan exhibit close phonetic similarity to their counterparts in HK, which have been split into high and low variants. The structure in (87) below is presented in such a way that the historical relationships between categories in Zhongshan and their correlates in HK are transparent. In terms of phonetic content, however, it is likely that Zhongshan has reanalyzed one of either [β] or [γ]:

(87) Zhongshan, assuming [α] > [β], [γ] HK Cantonese




While it is certainly possible to map Zhongshan features in (16) as [α] being a height contrast, [β] being a flat/rising contrast, and [γ] being a flat/falling contrast, there is a certain amount of redundancy to having two features, [β] and [γ], performing the same work, namely, differentiating flat from contoured tones. In addition, while we know that [γ] at some point in the past denoted a height contrast, and there is still a trace of this height distinction in [γ], evidenced by the fact that the formerly [high] tone Ia is high and flat and the formerly [low] tone Ib is falling, height does not seem to be the primary organizing principle in this domain, whereas height is clearly the contrast being made by [α]. Adopting the proposal that systems may stipulate directionality of contours via an already existing height contrast, with low tones being rising and high tones being falling by default (Barrie 2007; Yip 2001), we can postulate a reanalysis in Zhongshan of the [γ] feature as [β], here denoting [contour]. It is relevant to point out that this type of reanalysis does not entail any structural changes to our tree, as the relationships between categories stays constant. This type of reanalysis is also precisely what we would predict for a system attempting to maximize the economy of features (Martinet 1955; Clements 2003), or minimize complexity. This is discussed in more detail in §5.

(88) Zhongshan, [high] > [contour]




### 4.3.3 **Binyang**

Binyang county is a region approximately 200 kilometres west of Cenxi. Like Cenxi, Hong Kong and Guangzhou, the Binyang language has undergone a register split of multiple MC categories.

(89) Binyang (adapted from Tsuji 1977, Huang 2009)

| | | | | |
|---|---|---|---|---|
| i. | *kuei* | mid-rising | / 25 / | 'turtle' |
| ii. | *bei* | low-rising | / 13 / | 'surround' |
| iii. | *cou* | high-flat | / 44 / | 'nine' |
| iv. | *nui* | low-flat | / 22 / | 'woman' |
| v. | *mui* | high-falling | / 52 / | 'younger sister' |

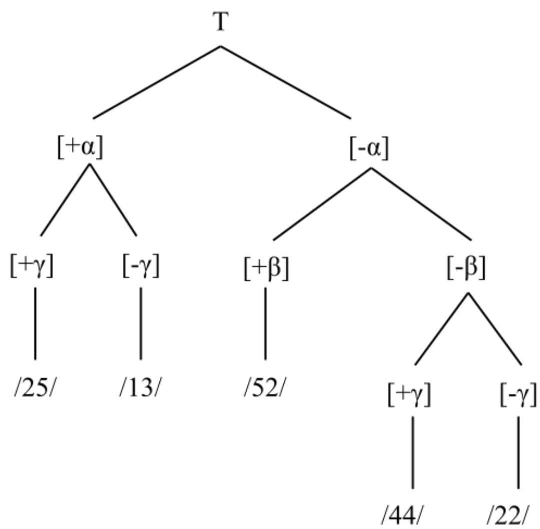
Binyang retains the tense-lax distinction in formerly TIV syllables, with the distribution of checked tones being analogous to HK/Guangzhou Cantonese. Category correspondences with MC are below:

| MC Onset | MC TI | MC TII | MC TIII | MC TIV |
|---|---|---|---|---|
| voiceless | /25/ (Ia) | **/44/** | /**52**/ | /5q, 3q/ |
| voiced | /13/ (Ib) | **/22/** | /**52**/ | /1q/ |

Table 20: Binyang (adapted from Lee 1993; Tsuji 1977, 1980)

Whereas Hong Kong, Guanghzou, and Cenxi exhibit a split in MC TI, TII, and TIII, Binyang appears only to have undergone a split in the TI and TII categories. Although Lee (1993), following McCoy (1966), speculates that there may have been a split in all three categories followed by a merger in TIII in Binyang, there is no evidence at the present to suggest one scenario over the other. The decision amounts to deciding whether we postulate a gradual loss of initial voicing contrasts coinciding with the introduction of tonal contrasts in Proto-Yue, with bifurcation into different groups of Yue languages at this time, or, a more rapid loss of Proto-Yue initial voicing contrasts and introduction of tonal contrasts, with bifurcations occurring afterwards. The present study has little to say on this matter. A contrastive hierarchy for Binyang is given in (90) below:

(90) Binyang [α] > [β] > [γ]




The clearest phonetic labelling of features in Binyang is [α] denoting a rising/not rising contrast, [β] denoting a falling/not falling contrast, and [γ] denoting a high/low contrast. Despite our [γ] feature continuing to denote a height contrast in modern Binyang, there is very clear evidence of phonetic drift in modern Binyang tones. The modern correlates of MC TII in Binyang, /44/ and /22/, are both flat, whereas they are rising in HK, Guangzhou, Cenxi, and Zhongshan. The modern correlates of MC TI in Binyang, /25/ and /13/ are both rising, whereas in all other languages thus far investigated they are flat or falling, and the modern correlate of MC TIII in Binyang, /52/ is falling, whereas the correlates of this category are flat in all other Yue languages investigated thus far. What is noteworthy here however is that despite there being no structural changes that are specific to Binyang, the structure of Binyang is **still** phonetically transparent with respect to historical categories after this drift occurred. This suggests that just as drift can motivate contrastive specifications, that contrastive structure can also constrain drift, especially in the absence of sandhi processes. The drift of a single tone therefore, rather than precipitate changes in the relationships between categories, may instead cause the synchronized drift of other tones to preserve structure.

### 4.3.4 Taishan

Taishan is the major dialect spoken in the Siyi (four districts) region, located in Guangdong province, southwest of Guangzhou. This region has historically been the source of much Chinese emigration in the late 19th and early 20th centuries, and was the source of virtually all Chinese immigration to North America during this time (McCoy 1966). The Siyi group of subdialects are largely "mutually unintelligible" (Wong 1982:72) with standard Cantonese. Although this group of dialects is, like Standard Cantonese, likely descendent from Middle Cantonese (78) and not simply Proto-Yue (77), a number of tonal developments render them distinct in several ways. We investigate standard Taishan (spoken in Taishan city) as well as a subdialect, Doushan Taishan, described in Wong (1982), spoken in Doushan commune.

(91) Taishan (adapted from Yue-Hashimoto 1972, 2005; McCoy 1966; Wong 1982)

| | | | | | |
|---|---|---|---|---|---|
| i. | *ji* | high-flat | / 55 / | 'chair' | |
| ii. | *ji* | mid-flat | / 33 / | 'clothes' | |
| iii. | *ji* | low-flat | / 11 / | 'to move' | |
| iv. | *ji* | low-falling | / 21 / | 'to use' | |
| v. | *ji* | mid-falling | / 32 / | 'different' | (Doushan /21/) |
| | | | | | |
| vi. | *jit* | high-checked | / 5q / | 'one' | |
| vii. | *het* | mid-checked | / 3q / | 'iron' | |
| viii. | *jit* | falling-checked | / 32q / | 'leisure' | (Doushan /21q/ ) |

Standard Taishan and Doushan Taishan are tonally quite distinct from Standard Cantonese, deviating both in terms of phonetic substance of inventory members, number of tones, and structural relations between tones. Most noteworthy is the fact that TIa and TIIIa, which share only a [γ] feature in Standard Cantonese, are merged completely in both Taishan dialects. While Standard Taishan and Doushan Taishan are themselves tonally quite similar to each other, they are distinguished by the presence of distinct TIIIb (/32/) and TIIb (/21) categories in Standard Taishan which have merged completely into a single category in Doushan Taishan (/21/) as below: [68]

| MC Onset | MC TI | MC TII | MC TIII | MC TIV |
|---|---|---|---|---|
| voiceless | /**33**/ (Ia) | /55/ | /**33**/ | /5q, 3q/ |
| voiced | /11/ (Ib) | /21/ | /32/ | /32q/ |

Table 21: Standard Taishan (Yue-Hashimoto 1972, 2005; McCoy 1966, adapted)

| MC Onset | MC TI | MC TII | MC TIII | MC TIV |
|---|---|---|---|---|
| voiceless | /**33**/ (Ia) | /55/ | /**33**/ | /5q, 3q/ |
| voiced | /11/ (Ib) | /**21**/ | /**21**/ | /21/ |

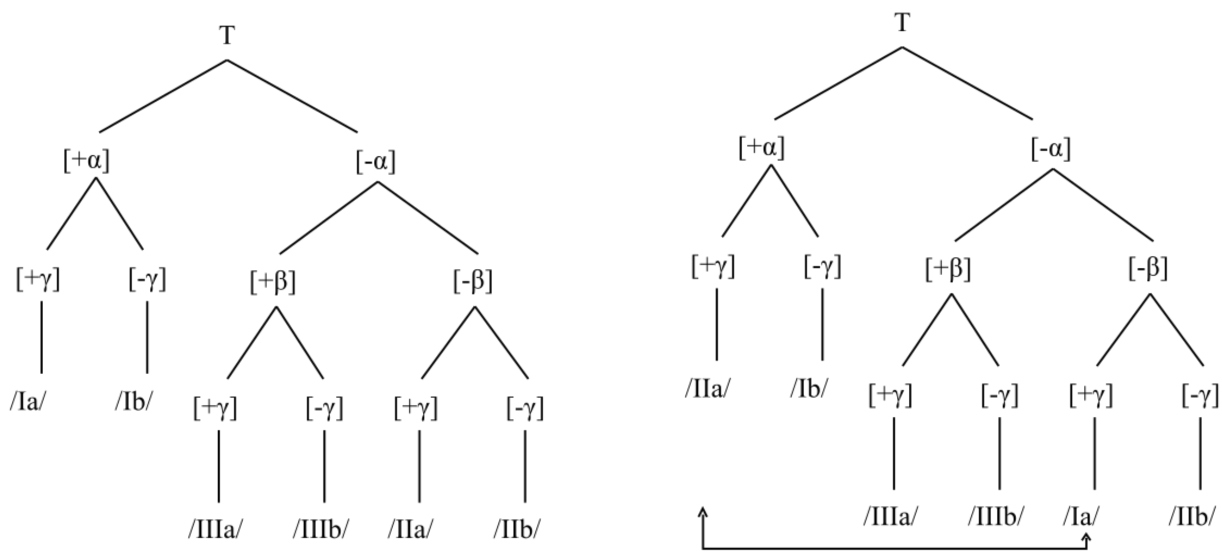
Table 22: Doushan Taishan (Yue-Hashimoto 1972, 2005; McCoy 1966; Wong 1982, adapted)

The Doushan TIIb/TIIIb merger and the general Taishan TIa/TIIIa merger provide information suggesting a fairly unique diachronic development for Taishan. Assuming, given the

68 Although the convention used by Yue-Hashimoto (2005) is to utilize a five-way numerical distinction to indicate tone values, she remarks that the low falling tone, /21/, is generally uttered so low as to 'give a cracking sound', and may sometimes co-occur with a syllable-final glottal stop.

*Sisterhood Merger Hypothesis*, that structural mergers occur only between contrastive sisters, we need to postulate a minimal sequence of changes which rendered both the TIIIa and TIa categories and the TIIb and TIIIb categories contrastive sisters. This requirement suggests an initial reanalysis whereby the Proto-Yue TIa and TIIa swapped respective positions on the tree, i.e., TIIa → [+α] and TIa → [-α]. [69]

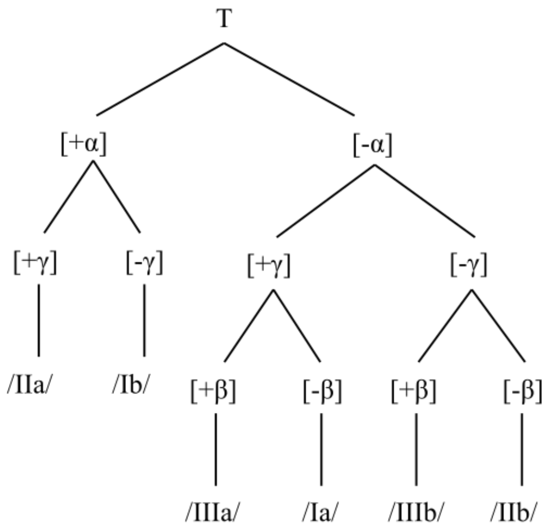
(92) Proto Yue → Proto-Taishan stage 1




Following the change exemplified in (92), it is evident that there was a reranking of features in the domain of [-α], such that [γ] > [β], as below:

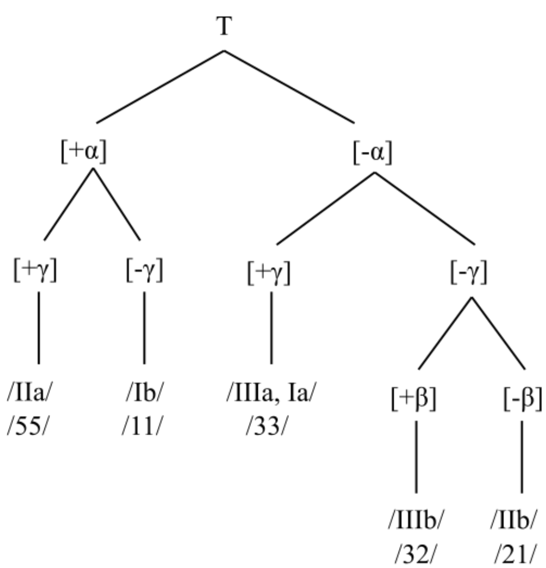
[69] Since this change is in a proto-language, it is somewhat difficult to precisely motivate this change. However, it is worth emphasizing the fact that in modern Taishan, IIa and Ia/IIIa are all flat, non-low tones. It is likely that [-β] did not make a substantial contribution to phonetic content in Proto-Taishan stage 1.

(93) Proto Taishan stage 2 (rerank)

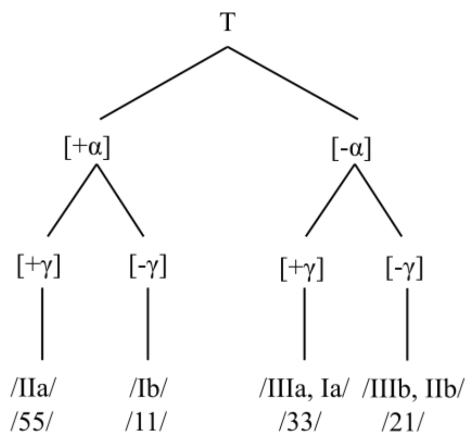



This yields a hierarchy where both TIIIa and TIa as well as TIIIb and TIIb are contrastive sisters, permitting both the merge of TIIIa and TIa, common to both languages, yielding (94) below, as well as the subsequent merge of TIIIb and TIIb in Doushan Taishan in (95) below.

(94) Standard Taishan (categories and values), Proto-Doushan Taishan (categories only)



(95) Doushan Taishan

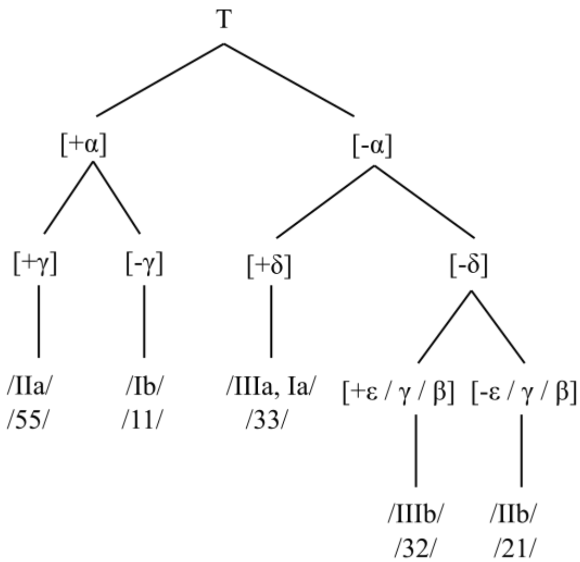



It is a good result for our theory that the changes postulated above for Taishan, motivated simply by what we know about the contrastive organization of Middle Chinese and the modern categories of Taishan, yield more or less phonetically sensible structures. The issue of phonetic labelling, however, brings up a fairly interesting problem for these two languages. Although the sequence of phonological changes outlined above suggests an [α] > [γ] > [β] type organization, the phonetic substance of both Taishan languages suggests that at some point in the past, [γ], which in Middle Chinese/Proto-Yue indicated a height contrast, in the domain of [-α] was subsequently reanalyzed as a contour contrast. This likely occurred before the Doushan Taishan TIIIb/TIIb merge which has been fully realized in the Proto-Doushan Taishan hierarchy in (95). I assume that the slightly falling contour reported for modern Doushan Taishan, /21/, is simply contrastively low, and thus [α] in (95) is a peripheral contrast and [γ] is a height contrast.

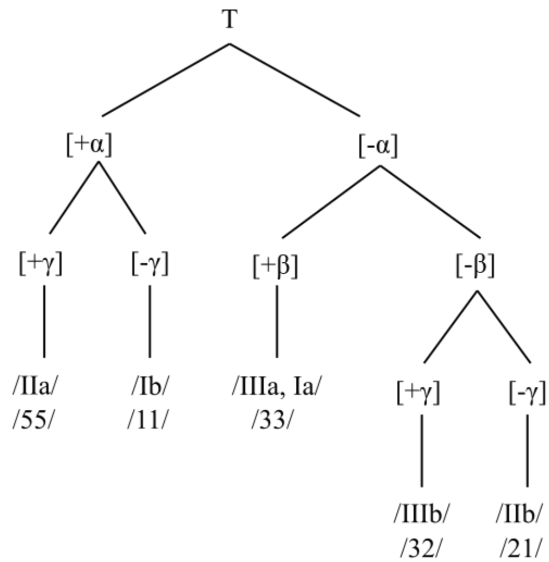
Also worth noting is that the [β] feature, which in (94) is the last vestigial remnant of the TIII/TII contrast from at least as far back as Early Middle Chinese, contrastively divides TIIIb and

TIIb on the basis of height in Standard Taishan (and presumably did so for Proto-Doushan Taishan).

(96) Standard Taishan (possible features)




Standard Taishan (revised features)




Regarding possible features, [α] serves as a peripheral/central distinction, [γ] (or [ε]) serves as a height distinction, and [δ] serves as a contour distinction. Regarding specific feature identities in Standard Taishan, we might postulate, since the most subordinate branching of Standard Taishan divides on the basis of height, that the division is **analogous** to the already existing height division in the [+α] domain, and thus utilizes a [γ] feature. One might also postulate a distinct feature due to reanalysis (e.g. [ε]). Finally, one may wish to emphasize the **homologous** origins of this structure present in Taishan and other related languages by utilizing a [β] feature here, underscoring the shared ancestry with the [β] division in Early Middle Chinese. All things considered, although it comes at the cost of obscuring some of the diachronic origins of the category splits, we choose a simple [α] > [β] > [γ] representation.

The fact that features may change their realization within a particular domain without changing consistently in all domains underscores the ontology of individual tonal identities as phonological objects. While phonetic naturalness likely guides diachronic change, the specifics of the diachronic sequence of changes for Taishan suggests that the determination of precise identities of features within a hierarchy is post-hoc relative to the acquisition of tones themselves. That is, features and hierarchies, rather than being the computational determiners of inventory content and phonetic substance, are **organizing principles** which influence and constrain acquisition and language change.

If the reverse were true, and universal features with static content were the sole determinants of tonal representation, it would take a fairly implausible sequence of changes to account for many of the developments outlined above. Assuming universal features, consider the case of the development of Standard Taishan in examples (92-94) above. Proto-Yue (92) is a 6-tone system. Any theory needs to postulate the existence of (at least) three phonologically active features, and we know that the most subordinate of these, [γ], divides on the basis of height. Regardless of the specifics of our universalist theory, [γ] is a phonologically real object. Fast forwarding to Standard Taishan, (94), we can observe that [γ] still divides on the basis of height in the domain of [+α], but does not divide on the basis of height in [-α]. How to contend with this issue? We can argue that [γ] in the domain of [-α] was reanalyzed at some point in the past, but it is unclear what tools a theory of universal features has to allow this. If features criterially determine underlying representation, and [γ] is a single object, in order to reanalyze only [γ] in the domain of [-α], this single phonological object had to, at some point prior, divide into two phonological objects with unique identities but which perform the same work: a [high] contrast in the domain of [+α], and another contrast, perhaps [low] or [high2] in the domain of [-α], after which a reanalysis of the [γ] feature in [-α] can occur. This division has to occur with no phonological consequence if we are to maintain the very strong likelihood that phonetic change is more or less gradual. If this did occur however, in addition to being utterly redundant, it is unclear what would have motivated this division in the first place, and the entire sequence becomes implausible.

### 4.3.5 **Yangjiang**

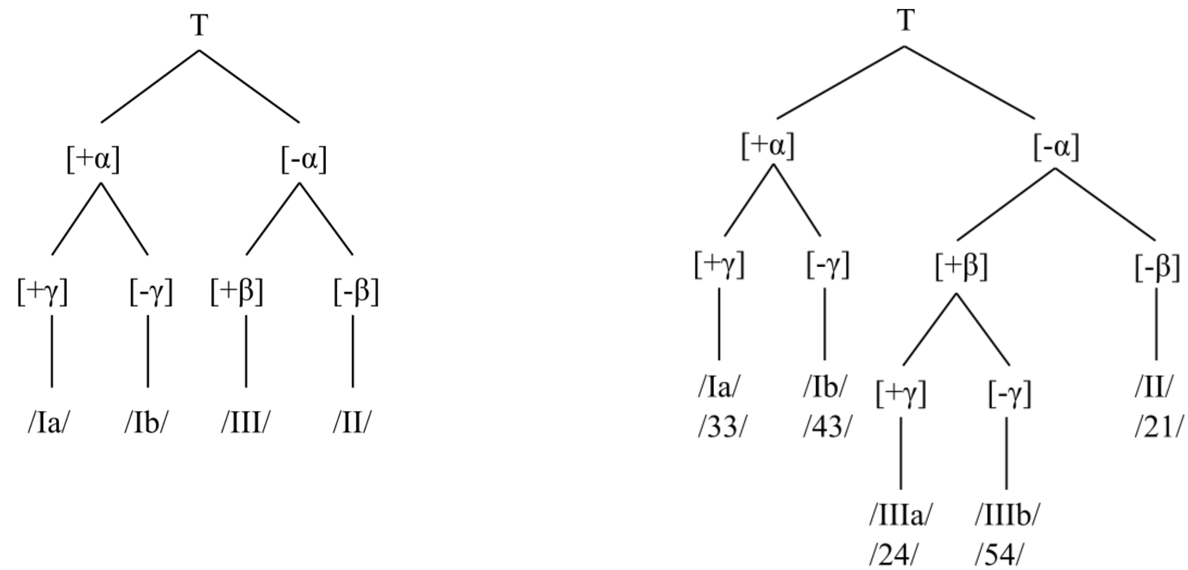
Yangjiang is one of two subdialects of the Gaoyang group of Yue languages, spoken in southwestern Guangdong. It is tonally distinguished from other Yue languages by the presence of a TIII and TI register split, but not a TII split, as below:

| MC Onset | MC TI | MC TII | MC TIII | MC TIV |
|---|---|---|---|---|
| voiceless | /33/ (Ia) | /21/ | /24/ | /24q, 21q/ |
| voiced | /43/ (Ib) | /21/ | /54/ | /54q/ |

Table 23: Yangjiang (adapted from Huang 1986:65; Lee 1993)

There is no evidence which suggests the existence of reanalyses in Yangjiang prior to the simple expansions in TI and TIII at the most subordinate level of the Late Middle Chinese tree, as below:

(97) Late Middle Chinese → (Proto-) Yangjiang




One interesting aspect of Yangjiang is the somewhat anomalous form that the modern phonetic content of its respective categories takes vis-à-vis their historical correlates. While [γ] in the tree above is, as in other Yue languages, the result of a register split due to the devoicing of voiced initials, the respective height levels of the resultant modern tones is the opposite of what we would normally expect given their values of [γ]. The modern correlates of the tones occurring on syllables with formerly voiced initials (/Ib/, /IIIb/) are all high rather than being low, and those occurring on formerly voiceless initials (/Ia/, /IIIa/) are all low rather than high. This raises an interesting problem: if the formerly low tones are now high, and the formerly high tones are now low, there must have existed in the past a point where the heights were equal, but at which point neutralization and category merges did not occur. We can account for this by reconstructing the precise path that these tones took, noting the fact that the formerly voiced initials are now both falling, in addition to being high. This implies the likelihood that the [γ] contrast changed at some point from denoting a height contrast to incorporating a contour contrast, resulting in a 'push' chain shift (Martinet 1952)[70] enabling drift past the point of equal tone height:

| | LMC/Proto YJ1 | | ProtoYJ2 | | ProtoYJ3 | | Yangjiang |
|---|---|---|---|---|---|---|---|
| [+γ] | high | → | high | → | low/mid | → | low |
| [-γ] | low | → | (mid) falling | → | (mid/high) falling | → | high-falling |

Table 24: Likely trajectory of [γ], Late Middle Chinese to Modern Yangjiang.

As with Binyang, the contrastive relationship between categories remained constant despite fairly drastic drift, supporting the idea that contrastive hierarchies coordinate and constrain phonetic drift of tones based on structural links and oppositions. In a theory without contrastive features, it is difficult to account for the fact that the contrastively low tones ([-γ]) ended up both high and falling, and the contrastively high tones ([+γ]) ended up low. The presence of a contrastive hierarchy also neatly accounts for how Yangjiang and Binyang end up with phonetic inventories which still transparently reflect the contrastive structure of Late Middle Chinese despite evidence of massive drift.

While we can be reasonably certain that the tree in (97) represents, at least, a Proto-Yangjiang feature hierarchy at some point after bifurcation from Middle Chinese and after the tonal split in the MC TIII category, constructing a precise hierarchy for modern Yangjiang is difficult without richer primary data, and we cannot be certain whether or not (97) represents the

70 It is unclear which tones or values pushed and which were pushed.

structure of modern Yangjiang. To start with, as with most other Yue languages, there is no data from activity which would otherwise serve to constrain the possibilities. Yangjiang is a relatively under-described Yue language, with modern sources citing Huang (1986).[71]

Chen (2000:17-19) aptly discusses interpretation of phonetic data in situations such as these:

> "with rare exceptions, phonetic transcriptions are based on aural judgment, and vary according to different practices and implicit assumptions on the part of the fieldworkers. Furthermore, the five-point pitch scale specifies a far greater number of tone shapes than one would ever need to describe any one language, as a consequence, forcing arbitrary choices upon the fieldworker in many cases. Take the tone shape [54] (attested in 57 dialects). One cannot tell a priori whether it is basically a high level tone [55] with a slight declination effect, or a variant of [53], or for that matter [454], and so forth. By the same token, if a dialect has only one rising tone, whether one transcribes it as [24], [34], [35], etc. depends as much on personal preferences and (implicit) theoretical assumptions as on the objective phonetic reality."

Chen provides the example of citation tones in Jianyang, a northern Min dialect spoken in Fujian province. Norman (1973) gives the following citation tones:

---

71 Bauer and Benedict (1997) cite Huang (1986). Lee (1993) cites Beijing Daxue (Peking University) for the source of her tonal shapes, but the reference is missing in the bibliography. Norman (1988) does not provide a reference for his data on Yangjiang, however, it accords with Lee (1993). Chen (2000) briefly mentions tonal categories in Yangjiang but does not provide a reference. Yue-Hashimoto (1972) cites older studies of Yangjiang syntax by Bo-Rong Huang (1958, 1959, 1966) but does not discuss tonal inventories in Yangjiang in any detail.

(98) Jianyang citation tones (adapted from Norman 1973, Chen 2000)

| **Even** | **Rising** | **Falling** |
|---|---|---|
| [33] | [35] | [53] |
| | | [43] |
| | | [32] |
| | | [31] |
| | | [21] |

Chen argues that this distribution of citation tones is unusually skewed. Not only is there a fairly marked tendency towards contoured tones, but this arrangement would be impossible under a classificatory system such as in Yip (1980, 1989). Instead, Chen argues, following Yip (1980:206), that three of the above falling tones, [43], [32], and [21], ought to be considered level tones ([44], [22], [11]) which exhibit downdrift or final-fall, which is a well-attested phonetic effect in tonological and intonational literature (e.g. Pierrehumbert 1980; Liberman and Pierrehumbert 1984; Tseng 1981; Shih 1988, 1991). We can thus revise the system as follows:

(99) Revised Jianyang citation tones (adapted from Chen 2000)

| **Even** | **Rising** | **Falling** |
|---|---|---|
| [44] | [35] | [53] |
| [33] | | |
| [22] | | [31] |
| [11] | | |

In interpreting phonetic data, it is thus important to keep the overall sound pattern in mind, and general heuristic principles, such as those discussed by Maddieson (1978), can be deployed to discount non-distinctive phonetic details.

This in mind, determining precisely the modern contrastive status of tones in Yangjiang becomes somewhat difficult without a richer data set. While it is noteworthy that both /43/ and /54/ are phonetically high and falling with respect to their respective counterparts which originated from the same tonal category (i.e. /33/ and /24/ respectively), whether [43] should be interpreted as /53/, /44/, /42/ etc., whether [54] should be interpreted as /53/, /55/ or /42/ etc., and whether [21] should be interpreted as /31/ or /11/, all have profound effects on the way the hierarchy is constructed by learners, especially in the absence of phonological activity. Nearly all combinations result in completely plausible systems (most of which are systemically more plausible than Huang's (1986) and Lee's (1993) reported values[72]), so I decline to speculate in the absence of more data.

## 4.4 Summary

This chapter motivated contrastive hierarchies for seven Yue languages. Unlike the Mandarin languages, these were based largely on phonetic inventory content mirroring the task that learners face when constructing phonological hypotheses in systems without activity (except in the case of Guangzhou). The phonetic data was coupled with historical data regarding the structure of Middle Chinese, and various mergers, splits, and reanalyses. Patterns of contrast

---

[72] To be fair, neither Huang (1986) nor Lee (1993) attempt to construct a phonologically plausible system for Yangjiang.

suggest that, along with contrastive hierarchies, complexity minimization, the subject of the next chapter, also serves as an organizing principle of phonologies.

# 5

# Complexity and phonetic naturalness in tonal systems

## 5.1 Overview

In the preceding chapters I analyzed tonal systems in several languages from two Chinese language groups, and showed that the contrastive hierarchy can be successfully applied to tonal systems, providing significant insight into their synchronic and diachronic behavior. Throughout the analysis I noted two recurrent patterns: i) that structural reanalysis, when we have sufficient evidence regarding the context in which it occurs, seems to be motivated by phonetic naturalness and feature economy; and ii) that languages with a greater number of phonological processes trend towards exhibiting less phonetically natural constellations of tones. In this chapter I make the argument that both of these patterns derive from a well-described, general tendency of the cognitive apparatus to minimize complexity in its representations, and that this can be formalized via advances in the study of information theory and generative geometry.

The following sections discuss and apply a theory of complexity to the findings in previous chapters. In §5.2 I discuss complexity and pattern recognition in cognition and information theory. In §5.3 I sketch the outlines of a theoretical framework towards applying these notions to tonal systems. I then show in §5.4 that reanalysis in the Mandarin languages; and in §5.5 that the overall structure, shape of inventory members, and reanalyses in Cantonese languages are all motivated by complexity considerations. A brief summary is provided in §5.6.

## 5.2 Complexity, Simplicity, and Pattern Recognition

Although it is both intuitive and uncontroversial that some continuous process of pattern recognition underlies much of cognition, determining the precise form or specific coding that these patterns take is a much more difficult and subtle matter. The ability of an individual to recognize, for example, a given ostinato or melodic figure in music (and consequently make predictions of how a melody might progress from some given arbitrary point within it), based on prior encountered empirical instances of similar figures, despite variations, differences in timbre, changing harmonies or keys and so on, necessitates the encoding of that musical pattern in some abstract form. Chater and Vitányi (2003:19) argue that we might postulate that an overarching cognitive process (which enables pattern recognition) and indeed an overall goal of the entire cognitive system is data compression: the encoding of data into such forms that can be recovered by some computable process.

But what forms do these patterns take? An issue that any cognitive theory which accounts for pattern recognition and data compression needs to contend with is the **problem of induction**: "for any finite body of observational data, there will be infinitely many patterns or regularities to which they all conform" (Boyd 1985:210). Chater and Vitányi (2003) illustrate this in a number of cognitive domains, but argue that certain candidate patterns are more cognitively plausible than others in a given data set:

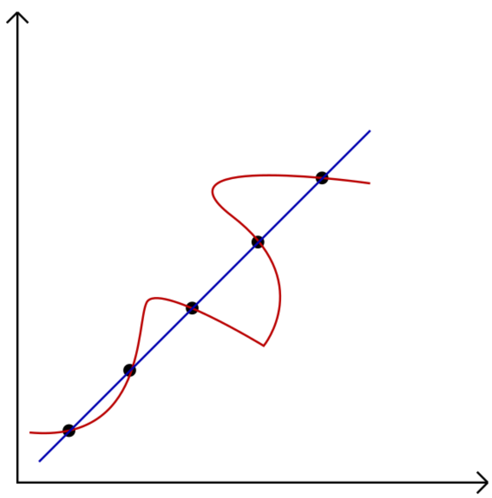


Figure 4: Two possible functions
(adapted from Chater and Vitányi 2003)

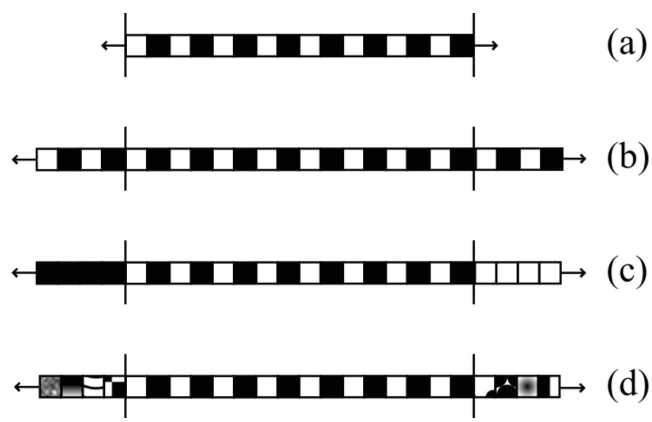



Figure 5: Three possible sequences
(adapted from Chater 1999)

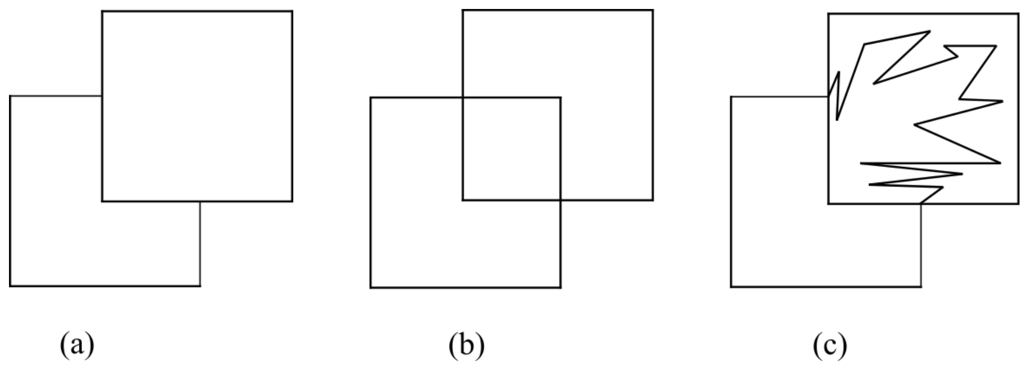



Figure 6: Two possible completions of an occluded square (adapted from Chater 1999)

In Figure 4, for example, there are an infinite number of continuous functions which pass through the points indicated on the plane; however, given only the points, all things being equal and in the absence of other considerations, a straight line passing through them would be hypothesized as the underlying function before the irregular one indicated. In Figure 5, despite the existence of other logically possible continuations of the sequence in (a) both to the left and to the right, the continuation in (b) would be postulated before other alternatives. As well in Figure 6, there is a similarly infinite number of possible completions of the partially occluded shape in (a); however, (b) seems intuitively more likely than (c). In each case only one logical possibility is seriously entertained despite there being an infinite number of logical alternatives, and it makes intuitive sense to conclude that these possibilities are in some way more 'cognitively natural' (Chater and Vitányi 2003:19) than others. Chater (1999:274) suggests that the cognitive system identifies patterns that are the *simplest possible* explanations for the available data, pointing out that there is a long history in the study of cognition of identifying simplicity.

Ernst Mach was an early proponent of the notion that the cognitive apparatus favours patterns that fit the simplest descriptions of data. Mach's oft-cited lecture on the economical nature of physical inquiry states: “when the human mind, with its limited powers, attempts to mirror in itself the rich life of the world… it has every reason for proceeding economically” (Mach 1895:186). Simple representations to Mach have the advantage of “disburdening the memory”, as,

> …in reality, the law always contains less than the fact itself, because it does not reproduce the fact as a whole but only in that aspect of it which is important for us, the rest being either intentionally or of necessity omitted. (1895:193)

The Gestalt tradition of perception too, describes the tendency for perceptual organization to maximize *prägnanz*, or pithiness, in creating the simplest and most orderly representations of perceptual data. Koffka (1935, 1962), a principal architect of this school of thought, proposed that the law of maximizing *prägnanz* underlies the grouping of perceptual data; the perceptual system reaches an equilibrium reflecting symmetric and simple representations when faced with otherwise chaotic data. Building upon this tradition, Hochberg and McAlister (1953) sought to quantify simplicity in information theoretic terms, arguing that the overall perceptual apparatus attempts to maximize simplicity in its representations: “the less the amount of information needed to define a given organization as compared to the other alternatives, the more likely that the figure will be so perceived” (1953:361). Peter van der Helm (to appear) makes a similar argument. He states that the stance of the simplicity principle, that “internal neuro-cognitive mechanisms tend to yield parsimonious percepts”, is “not only in line with Gestalt psychology, but is also sustained by… computational explanations” (to appear:16) and is a consequence of flexible cognitive architecture implemented in the relatively rigid neural architecture of the brain. Pothos and Chater (2002) identify the central role of simplicity in higher-level cognition. They conduct a series of experiments where participants were asked to, for example, categorize quasi-randomly placed dots into spatial clusters, place physical objects varying by inner and outer diameter into groups, and numerically rate the similarity of various pairs of stimuli.[73] They found that the simplest possible mathematical description of a grouping of stimuli predicts the categorization of those groups by participants. Research from a pure information theoretic approach finds that the simplicity of an encoding of a stimulus is negatively correlated with the perception of randomness, and that when individuals are unable to mentally encode a simpler representation of a stimulus, the 'experience of randomness' results (Falk and Konold 1997:301). Finally, Chater (1999:275) proposes that, as a guiding principle which goes beyond simply finding patterns in perceptual processing, identifying similarity, and scientific reasoning, simplicity has a more fundamental and "general role in cognition: ranging from reasoning to memory to learning" (1999:275).

As it turns out, simplicity, as conceptualized above, can be made mathematically precise and quantifiable as well. Kolmogorov complexity, or algorithmic entropy, is a measure of the amount of information required to describe a given object, *x* (Li and Vitányi 1997; see also Kolmogorov 1963, 1968). The Kolmogorov complexity of an object, denoted by $K(x)$, corresponds to the length of the shortest possible algorithm which outputs *x*. For example, Figure 7 below is a graphical representation of the Mandelbrot set. If the image were simply saved as an uncompressed 1600 by 900 pixel, 32-bit colour image file, encoding it would require approximately 46.08 million bits (=5.49 MB) of information. The shortest possible algorithm which can output this image,

---

73 The information structure of the experiments was kept constant. For example, the *x* and *y* coordinates of the respective dots in the first experiment corresponded to the inner and outer dimensions of the stimuli in the second. The rationale is that specific modality of information is of secondary importance to the abstract internal representations of data structure.

however, is a tiny fraction of this length, requiring only the quadratic recurrence equation for the Mandelbrot set and some small function for plotting the output graphically. It is worth pointing out that a small suitably implemented algorithm of this type would be able to generate an infinite number of image files of arbitrary size given enough computing power.

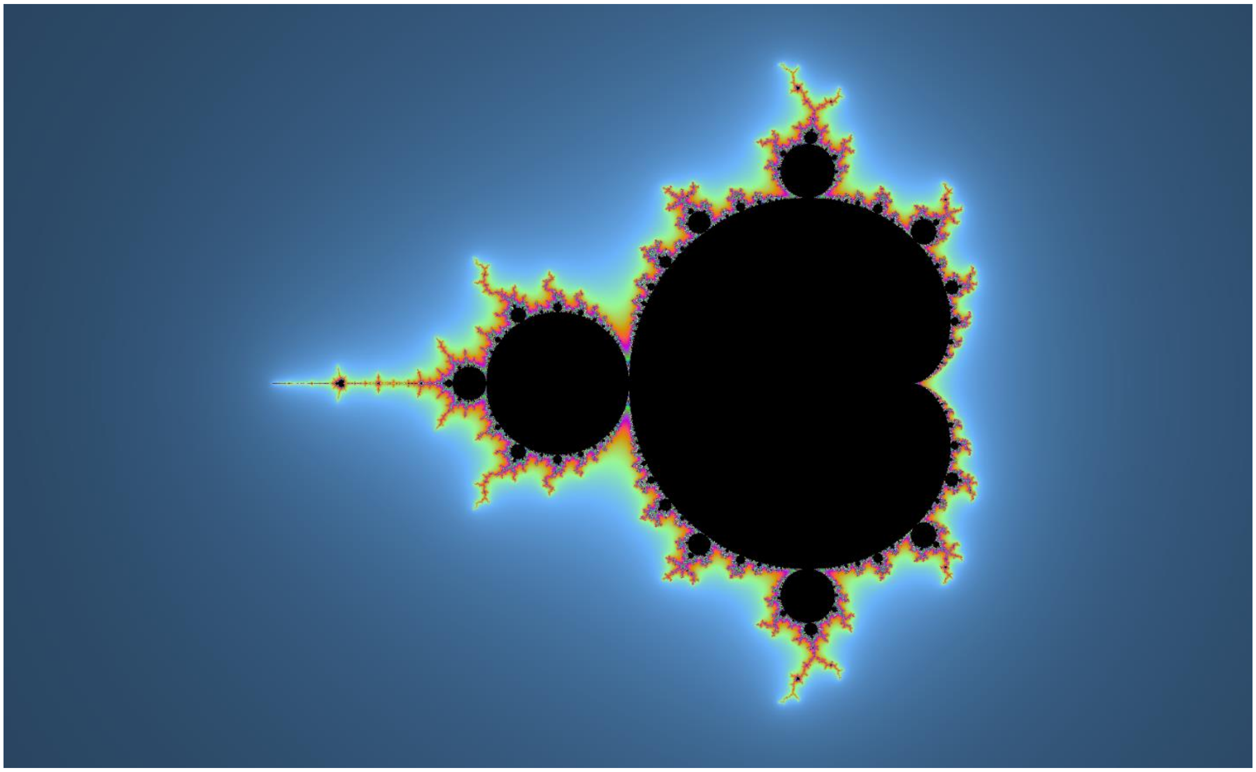


Figure 7: A graphical representation of the Mandelbrot set.

The specifics of the coding language turn out not to affect the overall complexity of the object to be described. The *invariance theorem* (Li and Vitányi 1997) states that *K(x)* is invariant and independent of the universal coding language it is written in, aside from some small additive constant. This is advantageous for our purposes because, as Chater (1999) points out, if we assume that the coding language is universal, we need not identify some detailed code that the cognitive system utilizes in specific; we can simply work at the level of algorithms. Chater and Vitányi argue that, cognitively speaking, pattern recognition in cognition can be understood to be closely related to the elimination of redundant information via compression (see also Attneave 1954; Barlow 1959), and, more importantly:

> formalizing simplicity provides a candidate solution to the problem of induction… the infinity of patterns, all compatible with any set of data, are not all equal: the cognitive system should prefer that pattern that gives the shortest code for the data. (Chater and Vitányi 2003:20)

Further evidence that the continuous calculation of probability and information content of objects is a basic cognitive process is found in Desalles' (2010; Saillenfest and Desalles 2012, amongst others) work on simplicity and unexpectedness. A discussion of Desalles' adaptation of Kolmogorov complexity and simplicity theory to measure unexpectedness will prove useful to our application of this theory to tonal systems.

I start with the observation that the probability, *P*, that a given Turing machine, *M*, outputs an object, *x*, given some random input can be approximated (Desalles 2010; Solomonoff 1978) as:

(100) $$P_M(x) \approx 2^{-K(x)}$$

where *K(x)* is the Kolmogorov complexity of *x*. Furthermore, the complexity of generating a given number, *n*, can be given as

(101) $$K_n \approx \log_2 n$$

This represents the fewest number of bits required to either instantiate a number, or locate it by rank in some ordered list. Imagine therefore a machine that converts strings of binary digits into letters according to the following table:

| | | | |
|---|---|---|---|
| *a* 000 | *b* 001 | *c* 010 | *d* 011 |
| *e* 100 | *f* 101 | *g* 110 | *h* 111 |

Table 25*:* A simple 3-bit character code table

Our object thus has a complexity of 3 bits, and the probability of our machine, given random input, outputting a string beginning with a given letter approaches 1/8.

That some objects have a lower complexity than others of the same length may appear counterintuitive at first glance. However, while the upper bound of complexity for a 7 digit number, for example, can be approximated by $\log_2(10^7)$, the shortest possible *algorithm* which outputs the number 7777777 simply consists of instantiating a single digit, *n*, and copying it *n* times, whereas a number such as 7829466 eludes such simple definition. In this case, the complexity of 7777777 can be represented as approximately the complexity of 7 plus the complexity of the copy operation (or the complexity of accessing the copy operation in our cognitive apparatus), whereas the complexity of 7829466 approaches the upper bound of complexity for a seven digit number.

Michael Leyton's generative theory of shape (2001) provides a useful starting point for formalizing some of the roles that complexity theory plays in cognitive processes. Leyton (2001:4) argues that the cognitive apparatus maximizes the reuse of structure:

(102) ***Maximization of Transfer**: Make one part of the generative sequence a transfer of another part of the generative sequence, whenever possible.*

To Leyton, the cognitive apparatus encodes the structure of objects by constructing them as multi-layered groups with their own internal structure. Each group is constructed recursively of a *fiber* group, which is a group of structure to be transferred, and a *control* group, which is the group of actions which transfers the *fiber* group. To illustrate, Leyton (2001:7) shows how human vision represents a square *generatively*, starting with a corner point, applying a translation function to trace a single side, then copying the structure by *transferring* the side as a *fiber* group, rotationally 90 degrees, and then repeating the rotational transfer process until a square results (Figure 8). Representing information generatively in this way both provides a simplest possible algorithmic encoding of the shape while accounting for an enormous number of psychological results (Leyton 1984, 1986a, 1986b, 1986c, 1992) that go well beyond visual organization, including psychological scales, categorization, grammatical structure etc. [74]

---

[74] Aesthetics, for example, notionally defined as the maximization of transfer and recoverability of structure, is given a rigorous mathematical description in Leyton (2001), with entire chapters devoted to architecture, mechanical design and music.

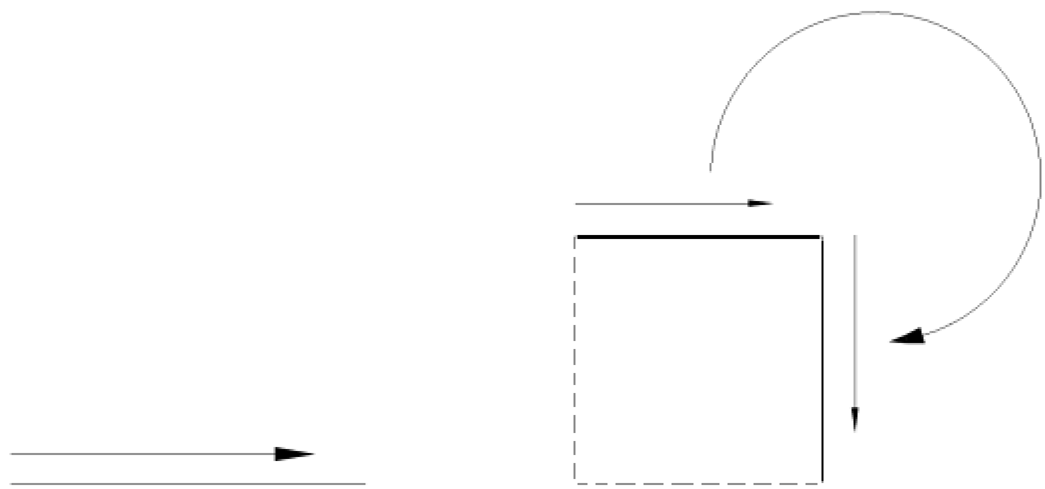


Figure 8: Generation of a side and transfer by rotation (adapted from Leyton 2001)

Desalles (2006) proposes that we can estimate the complexity of resultant structures, $C_R$, simply with the following general formula:

(103) $$C_R = C_F + C_T$$

Where $C_F$ represents the complexity of the *fiber* group and $C_T$ represents the complexity of the *transfer* group.

Consider then the following sequence, which does not seem random to any numerically literate human observer:

(104) 10 20 30 40 50 60 70

Desalles (2006) proposes a series of steps which yield the output in (104), noting that this is but one of multiple generative possibilities:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 1. | Start with an uninstantiated number: | - | | | | | | |
| 2. | duplicate via translation: | -- | | | | | | |
| 3. | instantiate the second digit: | -0 | | | | | | |
| 4. | transfer via translation: | -0 | -0 | -0 | -0 | -0 | -0 | -0 |
| 5. | dissociate the transfer operator | -0 | -0 | -0 | -0 | -0 | -0 | -0 |
| 6. | instantiate/increment the first digit | 10 | 20 | 30 | 40 | 50 | 60 | 70 |

The generation of this sequence maximizes transfer. Not only does each step maximize the reuse of structure in *fiber* groups, the operators themselves maximize transfer. The *duplicate* operator in step 2, for example, becomes a *copy* operator in 4, and a *copy* + *increment* operator in 6. We can thus calculate the complexity of the entire sequence as $C_{dup}$ for step 2; $C_0$ for the instantiation in step 3;[75] $C_{copy}\ (\approx C_{dup})$ for step 4; and $C_{copy}$ + $C_{inc}$ for steps 5 and 6. Since the generative sequence reuses operators in steps 2, 4 and 6, the total complexity of the sequence is proportionate to the complexity of the *copy* operator (plus some small constant), the *increment* operator, and the complexity of the digits 1 (the initial digit to be incremented) and the digit 0 (the second digit instantiated). If we wish to stipulate that the sequence is bounded at 7 elements, we add the complexity of 7. If the operators utilized in this sequence are already represented in the cognitive apparatus, we need not even include the complexity of the operators themselves to generate the sequence, simply the complexity of accessing them. Any person familiar with basic

[75] Desalles (2006) points out that the complexity of 0 is not necessarily zero. If for example 0 follows 9 in an ordered list, the complexity would instead be $\log_2 10 \approx 3.3$.

mathematics would be able to continue the sequence indefinitely, but would not require an infinite amount of information to do so.

Desalles (2006) reports on an experiment which shows that such calculations are done automatically all the time. Participants were offered a free lottery ticket, and had to choose two among fourteen combinations of six numbers ranging from 1 to 49. The number sequences varied by complexity, with some of the sequences with the lowest possible complexity being [1 2 3 4 5 6], [34 35 36 37 38 39] and [10 11 12 44 45 46]. Although all possible sequences are, of course, equally probable, participants showed a strong avoidance of simpler sequences, with the strength of avoidance correlating inversely with the complexity of the sequence. Typical elicited explanations for this were “if you play that one, it’s sure you won’t win”, or “it would be stupid to play that one”.[76] Less complex sequences were systematically rated as being less likely by participants, motivating the hypothesis that unexpectedness of an event *U(x)*, is expressible as the difference between the expected complexity, $C_{exp}(x)$ (which, in this example would be approximately $6C_{49} = 33.7$) , and the observed complexity $C_{obs}(x)$ (which, in the case of [1 2 3 4 5 6] was calculated to be approximately 3.[77,78]

---

76 Desalles notes that Savoie and Ladouceur (1995) also report avoidance of regular structures by lottery players.

77 Of some relevance is the observation that certain low complexity sequences were not avoided— [9 13 18 24 31 39] is one cited example, and this is attributed to the fact that the algorithm for generating the sequence (increment the gap between adjacent numbers by 1) may not be immediately recognizable to participants.

78 A full review of phenomena such as these is beyond the scope of this paper, however similar results concerning the relationship between the sense of unexpectedness and the difference between observed and expected complexity have been reported in diverse domains such as conversational interest in narratives (Desalles 2010), moral dilemmas (Saillenfest and Desalles 2012) and perception of coincidences Desalles (2008).

In the preceding section I provided a brief outline of complexity theory. I discussed numerous strands of evidence that, taken together, suggests that the cognitive apparatus seeks to minimize complexity in its representations, and that this preference is domain general. I also discussed research showing that complexity is formalizable and invariant, being independent of the coding language utilized. I argued that we can conceptualize formal objects as being algorithmically derived, and proposed that utilizing bundles of operators recursively applied in sequence to some structural primitive provides a means for describing the relative complexity of objects, including phonological systems.

## 5.3 Features and Complexity

Discussing the history of phonological theory, Dresher (2009:5) points out that there has been a recurring tension between two views situated on rather opposing poles on how members of phonological inventories should be defined. On one extreme is the view taken by Saussure (1972 [1916]: 166) that "in a language there are only differences, and no positive terms". In other words, from this perspective phonemes are defined in purely negative, abstract terms, where oppositions or contrasts are of sole importance. The other extreme is exemplified by research by Bloomfield (1933:79) who supposed that phonemes are definable via invariant acoustic cues "present in the sound waves". Dresher goes on to argue that the position of various schools of thought at various times have resolved the tension between these two poles in different ways, but that the Prague school offered a particularly balanced solution: "Jakobson (1941) emphasized the oppositional nature of phonemes; but these oppositions are made in terms of distinctive features that have substantive content… for example… Jakobson proposes that learners begin with an undifferentiated representation which first splits into a consonant (C) and a vowel (V). This formal opposition has phonetic content: V represents a sound with greater sonority, and C one of lesser sonority.” (2009:6).

A central principle in complexity theory, **Maximization of Transfer**, encapsulated in (102) above, which refers to the tendency of the cognitive apparatus to reuse structure wherever possible, is fundamentally compatible with feature theory in general, and especially with the conceptualization of features advanced here. Each phoneme in an inventory contains its own unique configuration of features, but the phonetic content contained in a given feature itself is replicated across the entire set of phonemes possessing some value of the feature in question. If features are therefore to be conceptualized as mappings of elements of phonetic content onto an abstract system of oppositions and interrelations, we can understand the ‘content’ of a given feature to be a function or set of functions, with the value of the feature serving as an input to that function. Each feature is therefore a bundle of structure, and the placement of an individual tone on a tree represents its derivational pathway. In the terminology of the discussion above, a *fiber group*, construable as a pitch primitive, is transferred recursively through successive *control groups,* with individual features representing individual operations or groups of operations to be performed.

As an example, let us examine a hypothetical set of steps a learner of Songjiang (a language spoken in Shanghai, example (1) in §2, reproduced below) might go through in acquiring the following tonal inventory. Note that in this particular language there is a correlation between syllable-initial voicing and tone height. We can ignore this technicality for the time being and assume that height is the primary contrast.[79]

(105) Tonal inventory of Songjiang

| | | | | |
|---|---|---|---|---|
| *yin ping* (high-flat) | Ia | 53 | *ti* | 'low' |
| *yin shang* (high-falling) | IIa | 44 | *ti* | 'bottom' |
| *yin qu* (high-rising) | IIIa | 35 | *ti* | 'emperor' |
| *yin ru* (high-short) | IVa | 5 | *paʔ* | 'hundred' |
| | | | | |
| *yang ping* (low-flat) | Ib | 31 | *di* | 'lift' |
| *yang shang* (low-falling) | IIb | 22 | *di* | 'brother' |
| *yang qu* (low-rising) | IIIb | 13 | *di* | 'field' |
| *yang ru* (low-short) | IVb | 3 | *baʔ* | 'white' |

First let us adopt the assumption that the Successive Division Algorithm (Dresher 1998), reproduced as (106) below, refers to an actual acquisitional process.

79 We assume the construction of this hierarchy based solely on tonal inventory.

(106) The Successive Division Algorithm

a. Begin with *no* feature specifications: assume all sounds are allophones of a single undifferentiated phoneme.

b. If the set is found to consist of more than one contrasting member, select a feature and divide the set into as many subsets ass the feature allows.

c. Repeat step (b) in each subset: keep dividing up the inventory into sets, applying successive features in turn, until every set has only one member.

Let us also assume that, in line with the convention established above for other languages descended from Middle Chinese, that so-called *short* or *checked* tones are allotonic with their respective full-tone counterparts. We are thus left with a 6-tone inventory, initially undifferentiated for our hypothetical learner, consisting of {/53/, /44/, /35/, /31/, /22/, /13/}. We start out with the fact that this particular inventory consists solely of six **tones**, implying that our learner has already differentiated tonemes from a set of segmental phonemes. Since tones, by definition, have some obligatory pitch, it is reasonable to assume at this stage of acquisition the presence of some point of pitch, our initial *fiber group,* as a structural primitive and starting point in differentiating our inventory from segmental phonemes. We thus denote this pitch primitive at arbitrary height, *n,* with no phonological domain of application simply as [n]. Since all tones in Songjiang are bound to syllables, we postulate some operation which **translates** a point of pitch across a given syllable, which can be conceptualized as an operator, *trans,* which takes a point of pitch and copies it across the relevant phonological domain. This operator is not associated with the contrastive status of any individual feature, being present in the undifferentiated set of phonemes, and can thus be thought of as bound to the entire tonal system, rather than to any feature. The application of *trans* to [n] thus results in the following:

(107) *trans* applied to a pitch primitive in an undifferentiated set

[**n** **n** **n**]$_\sigma$

We suppose, purely for the purpose of discussion, that the syllable is comprised of three length units, capturing the capability to distinguish flat from rising or falling and from concave or convex tones.[80]

Our learner might initially note that there is some sort of pitch height contrast:

(108) Songjiang, Stage I

| | /53/ | /44/ | /35/ | /31/ | /22/ | /13/ |
|---|---|---|---|---|---|---|
| [high] | + | + | + | - | - | - |

Stage I differs from the undifferentiated inventory by the presence of the [high] feature. Although the undifferentiated inventory possesses some arbitrary, obligatory height, our Stage I inventory differs in that the specific height of a tone now matters, and the cognitive apparatus now contains the encoding capacity to phonologically differentiate tones based on height. That height is thus specifically ***instantiated*** at some pitch which contrasts the two possible values. Let us call the

80 Discussion of length units here is purely illustrative, and serves only an expository purpose. There is no evidence here suggesting that syllables are comprised of *n* discrete temporal units.

operator which performs this *inst,* assuming, for the purposes of discussion that it instantiates high tones at average pitch (i.e. tone digit 3) plus 1 tone digit, and low tones at average pitch minus 1 tone digit.[81] We can represent this two tone system as follows:

(109) The application of *inst* to (107)

| | | | |
|---|---|---|---|
| [**4** | **4** | **4**]$_\sigma$ , | [+high] (n= +1) |
| [**2** | **2** | **2**]$_\sigma$ | [-high] (n= -1) |

Suppose then that our learner further subdivides the inventory, noting that some tones are flat, and some are contoured:

(110) Songjiang, Stage II

| | /53/ | /44/ | /35/ | /31/ | /22/ | /13/ |
|---|---|---|---|---|---|---|
| [high] | + | + | + | - | - | - |
| [contour] | + | - | + | + | - | + |

At this stage of the division, directionality of pitch change is not important, only the presence or absence of it. We can conceptualize this structurally as some incremental operator applied to *trans* which serves to ***shift*** the pitch in some direction across the syllable, *shift.* Since we wish to maximize transfer, we predict that *shift* borrows most of its structure from *trans.*

81 Again, this is for the purposes of illustration only and does not imply the absence of variation. One would expect variation up to the point of contrast.

(111) application of *shift* to (109)[82]

| | | | |
|---|---|---|---|
| [4-$\boldsymbol{i}$ | 4+$\boldsymbol{0i}$ | 4+$\boldsymbol{i}$]$_\sigma$ | [+contour +high], ($i \neq 0$) |
| [4 | 4 | 4]$_\sigma$ | [-contour +high] |
| | | | |
| [2-$\boldsymbol{i}$ | 2+$\boldsymbol{0i}$ | 2+$\boldsymbol{i}$]$_\sigma$ | [+contour –high], ($i \neq 0$) |
| [2 | 2 | 2]$_\sigma$ | [-contour –high] |

Finally our learner acquires a third feature, denoting directionality:

(112) Songjiang, Stage III

| | /53/ | /44/ | /35/ | /31/ | /22/ | /13/ |
|---|---|---|---|---|---|---|
| [high] | + | + | + | - | - | - |
| [contour] | + | - | + | + | - | + |
| [falling] | + | | - | + | | - |

The final feature, here arbitrarily denoted as [falling] (rather than [rising]), distinguishes the remaining tones in terms of pitch directionality. We might conceptualize this operator, *inv*, as inverting the direction of the of *shift* operator.

82 We conceptualize *shift* here as akin to a skew operation: lowering initial pitch and raising final pitch or vice versa, with pitch in the centre constant for contoured tones. Again, this is illustrative only and variation exists.

(113) *inv* applied to (111)

| | | | |
|---|---|---|---|
| [4-(**-1**) | 4+0(**-1**) | 4+(**-1**)]$_{\sigma}$ | [+contour +high +falling], *i* = -1 |
| [4-(**1**) | 4+0(**1**) | 4+(**1**)]$_{\sigma}$ | [+contour +high –falling], *i* = 1 |
| [4 | 4 | 4]$_{\sigma}$ | [-contour +high] |
| | | | |
| [2-(**-1**) | 2+0(**-1**) | 2+(**-1**)]$_{\sigma}$ | [+contour -high +falling], *i* = -1 |
| [2-(**1**) | 2+0(**1**) | 2+(**1**)]$_{\sigma}$ | [+contour -high –falling], *i* = 1 |
| [2 | 2 | 2]$_{\sigma}$ | [-contour -high] |

(114) resultant tones

| | | | | |
|---|---|---|---|---|
| [5 | 4 | 3]$_{\sigma}$ | [+contour +high +falling] | = /53/ |
| [3 | 4 | 5]$_{\sigma}$ | [+contour +high –falling] | = /35/ |
| [4 | 4 | 4]$_{\sigma}$ | [-contour +high] | = /44/ |
| | | | | |
| [3 | 2 | 1]$_{\sigma}$ | [+contour -high +falling] | = /31/ |
| [1 | 2 | 3]$_{\sigma}$ | [+contour -high –falling] | = /13/ |
| [2 | 2 | 2]$_{\sigma}$ | [-contour -high] | = /22/ |

It is worth underscoring that with all operators, detailed specifics of their encodings are potentially language- or even individual-specific.

The smallest possible encoding of each of our operators, *trans*, *inst*, *shift*, *inv*, is worth discussing. The first of these, *trans*, can be considered ‘free’, in the sense that it is obligatory in any tonal language. This does not imply that its actual complexity is zero, but that the relative increase in complexity of including it in our system is zero, since all systems must include it by default. The complexity of *inst* is in the same order of the complexity of a single binary switch or small constant. The operator *inv* is also likely of a similar size. The last operator here, *shift*, also has complexity similar to that of a small constant, plus the complexity of *trans*. If we conceptualize *trans* as taking a pitch primitive and copying it over the duration of a syllable, *shift* utilizes the same operation but applies a small increment to each stage of the translation. Different rates of change can be accounted for by different constants, changes in rate over time can be accounted for by applying *shift* recursively to its own output.

Interestingly, identifying the operators which generate phonetic content in this way allows us to compare, in a general manner, the total complexity of this type of algorithmic encoding of tonemes with the complexity of a hypothetical featureless system where the six tones in Songjiang are simply memorized, or stipulated, rather than being generated by features in the phonology. The total complexity of the system outlined above can be approximated as $C_{trans} + C_{shift} + C_{inst} + C_{inv}$ . Since the complexity *inst*, *inv* are comparable to the complexity of single digit binary numbers, and the complexity of *shift* is comparable to the complexity of *trans* plus some small constant, total complexity is in the order of $2C_{trans} + 2C_2 + C_n$, and likely not more than $2C_{trans} + 3C_n$. We would also need to include the complexity of the positive and negative values for the features themselves, each a single binary digit (=$5C_2$, there being *n*-1 slots for binary features in any *n* tone system). The complexity of six memorized tones is substantially higher. In fact, in order to postulate that tones are memorized at all, we would need to stipulate that absolute pitch for each tone is also memorized. The notion of a tone being *relatively* or *contrastively* higher or lower than another given tone is, by its very nature, algorithmic. Since the just-noticeable-difference for frequency discrimination in normal speech with 80Hz < F0 < 500Hz is approximately 1Hz (Kollmeier et al. 2008:65)[83], the total complexity of each of the 6 memorized tones would be approximately in the order of $C_{l\cdot(2m+1)^n}$ , where $l$ = (500-80), or the initial pitch of the tone, $n$ = the number of possible temporal divisions of a syllable where discrete instances of pitch are producible or perceptible, and $m$ = the maximum number of Hertz by which an individual is capable of changing the continued uttered frequency of their voice from one temporal division to the next.

Even this estimate ($C_{l\cdot(2m+1)^n}$) contains, implicitly, some element of compression, since it reduces the total number of possible combinations of frequencies in their respective temporal slots in normal speech range to only those which are capable of being uttered, therefore $C_{l^n}$ is a better upper bound. The total complexity of this system is thus 6 $C_{l^n}$, compared with $2C_{trans} + 5C_2 + 3C_n$, for the algorithmic component of our features. The complexity of our *trans* operator is bounded, and less than the complexity of a single flat tone instantiated at an arbitrary but specific pitch level. By definition, the complexity of an object is the length of the smallest possible algorithm required to generate it. An absolute upper bound for the complexity of *trans* thus equals the complexity of a single memorized flat tone at arbitrary pitch, since one possible (albeit highly inefficient) algorithm is simply to stipulate the tone.[84] We can therefore represent the extreme upper bound for the algorithmic generation of tones as $2C_{l^n} + 5C_2 + 3C_n$, and a more reasonable

---

83 Liu (2013) gives a range of 4-10Hz JDN for Mandarin speakers.

84 It is certainly lower than this, as one could save information simply by stipulating two variables for flat tones, the left edge height, *h*, and the length, *n*, rather than stipulating *n* instances of *h*.

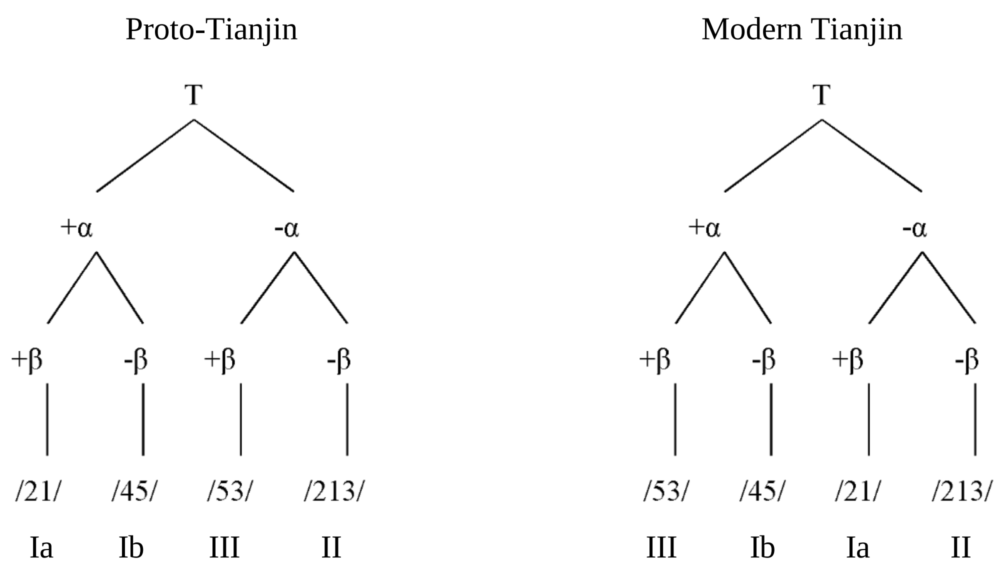
upper bound as $2C_{l \cdot n} + 5C_2 + \ 3C_n$ which are both less than $6\ C_{l^n}$ for all realistic values of $l$ and $n$.

The above comparison takes, of course, somewhat of an extreme, absurd position in calculating the complexity of stipulative memorized tones. There is absolutely no evidence that languages attend to specific pitch levels, and all of the evidence points to tones being differentiated contrastively rather than absolutely. It just so happens that contrastive differences are by definition both algorithmic and structurally minimal. The model implied by the discussion above inherently seeks to be as theoretically parsimonious as possible. One strength of this model is not only internal parsimony and bounded total complexity, but the capability to quantify and compare parsimony of theoretical alternatives. In addition, the discussion in §5.2 situates the concept of parsimony in the construction of theory itself into a broader cognitive and mathematical framework. Parsimony as an organizing principle of cognitive architecture is ubiquitous and theoretically quantifiable. Be that as it may, it is however important to point out that, for our purposes, we needn't ever actually calculate specific complexities for any of these systems. We only need to point out, at a minimum, the required operators, and, in some cases, their complexity relative to each other. In cases of reanalysis (or the lack thereof), for example, we need to compare the complexity of one hierarchical configuration with the complexity of a phonologically possible alternative, and the presence vs. the absence of operators in hypothetical systems permit this.  This is the subject of the following section.

## 5.4 Complexity in activity-rich languages: Tianjin and Nantong

Tianjin Mandarin provides a suitable starting point for a discussion on reanalysis. The reanalysis in Tianjin, where the modern correlates of Middle Chinese tones III and Ia swapped respective positions, is reproduced below:

(115) Contrastive reanalysis of Proto-Tianjin




I argued in §3.2.4 that it is was worthwhile asking why Tianjin Mandarin underwent a reanalysis whereas Beijing and Pingyao did not, since the Beijing and Pingyao inventories are at least as phonetically unnatural as that of Proto-Tianjin. We eventually concluded that neither of Beijing nor Pingyao represented any possible alternative systems which were **more** phonetically natural. Possible feature hierarchies in Beijing and Pingyao are heavily constrained by the abundance of sandhi processes, many of which are shared and almost certainly date back to at least Late Middle Chinese. Although Tianjin also exhibits an unusually high number of sandhi processes, it was shown in §3.2.4 that the two falling tone dissimilatory sandhi processes which lock Tianjin into its present configuration, (/21/→ 213 / __/21/ and /53/ → 45 /__/21/) in (36) and (37) are unique to Tianjin, being acquired after reanalysis. This implies the existence of a point in time where activity patterns permitted multiple feature hierarchies, since any hierarchy at this point in time only needed to associate /213/ and /45/, the two tones which enter into the famous 'third tone sandhi' (in (35)) process common to most Mandarin languages and still present in Modern Tianjin.[85]

Comparing the two trees, we can note that there is absolutely no change in what our [β] feature is doing from Proto-Tianjin to Modern Tianjin. /53/ and /21/ are both [+β], and /45/ and /213/ are both [-β] in each case. Put another way, this implies that the similarities and contrasts between inventory members invoked via both the values and content of [β] did not change during reanalysis. Rather, the reanalysis is centered solely on the [α] feature. While /45/ and /213/ keep their values as [+α] and [-α] respectively, /21/ changes from being [+α] to [-α] and /53/ changes from being [-α] to [+α].

This change is purely at the level of contrast. That is, while the individual tones remain phonetically constant at the point of reanalysis, the realization of the features themselves change in a way such that they output the same tones with a different algorithm. To begin with, while all tones in our undifferentiated set in Tianjin are, in the absence of specific contrastive features, the products of at least the *trans* operator, unlike Songjiang, Tianjin lacks any flat tones altogether.

---

85 It has very little consequence for the present analysis if the process in (38), /53/ → 21 / __ /53/, is, or was a robust sandhi process at any point. A phonological rule of this type associates /53/ and /21/, which, by virtue of the third tone sandhi rule in (35) would already share a feature in all symmetric, 2-feature systems. Asymmetric, 3-tone systems are unlikely in this context as we would need to postulate the symmetric 2-feature Middle Chinese transitioning into an asymmetric 3-feature Proto-Tianjin before reverting back into the modern 2-feature system. There is no evidence that this is the case.

Therefore in addition to *trans,* which translates a pitch primitive across a syllable, all tones are also necessarily subject to an increment operator, *shift*, shifting the pitch in an unspecified direction across the syllable. Minimality considerations dictate that *shift* apply to all undifferentiated tones, rather than being bound to multiple features (as this implies less structural cost), thus all tones in Tianjin are directional by default, with the height and directionality unspecified in the absence of features. In both Proto-Tianjin and modern Tianjin, [β] performs the same work, specifying directionality, with [+β] tones being falling and [-β] tones being rising. [β] in both Proto- and modern Tianjin contains, minimally, an *inv* operator consisting of a binary switch, which instantiates and inverts the directionality of our otherwise undifferentiated directional tones.

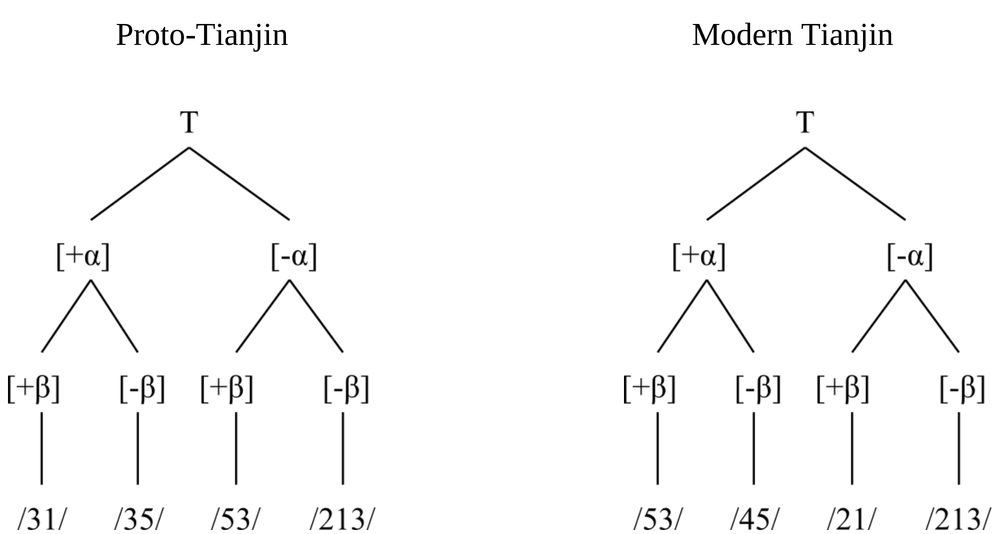
[α] in Proto-Tianjin is a [diffusing], or [peripheral] feature, contrasting those tones which move towards the periphery of the pitch range (that is, [+α]) with those which move towards the centre (that is [-α]). In modern Tianjin, [α] is a simpler high ([+α]) vs. low ([-α]) contrast. Features such as [diffusing] and [peripheral] are necessarily more complex than height features. As in Songjiang, all that is necessary to invoke height in modern Tianjin is a single *inst* operator, which instantiates the pitch at a specific level, with complexity in the order of a small constant. [α] in Proto-Tianjin however requires some extra work. While [α] here also must necessarily incorporate an *inst* operator, which divides the pitch into high and low ranges, there needs to be an additional operator, *check,* which identifies whether a tone is rising or falling: [+α] tones are low if falling and high if rising, [-α] tones are high if falling and low if rising. [86]

---

[86] This analysis assumes that the low concave tone is underlyingly low-rising. This predicts that speakers of Tianjin would perceive any low rising tone as the low concave tone. More robust acoustic data on specific contours, coupled with perceptual data would likely enable a more granular analysis of finer phonetic content.

Another possibility is that since we know that, compared to modern Tianjin values, Middle Chinese Ia was high and fell while Ib was low and rose, tonal values in Proto-Tianjin could have been subtly different than modern values. While the precise timing of the structural reanalysis vis-à-vis tonal drift is unknown, [+α] tones at one point were likely both of mid height, while [-α] tones at this point were both high and low. In this scenario, [+α] tones would have been central, and [-α] would have been peripheral, with further tonal drift precipitating a reanalysis into the modern [high] vs. [low] configuration, as illustrated below. If this was in fact the case, then [α] at this point in time is even more complex, involving the division of pitch into high, low, and mid, rather than just high and low, while also involving *check* to determine the height of [-α] tones. Thus in this scenario, [+α] tones would have been instantiated at mid via *inst*, [-α] tones would have been instantiated at mid plus or minus *n* tone digit increments, with the determination of whether *n* is positive or negative based on the directionality of a tone, via *check*, which returns a positive value if a tone is falling, and a negative one if a tone is rising.

(116) Alternative Proto-Tianjin reanalysis: [α] = [central/peripheral] → [high]



Since the complexity of the rest of the system is the same for both Proto- and modern Tianjin, the structural reanalysis to modern Tianjin represents a drop in overall complexity: $C = C_{default} + C_{\alpha} + C_{\beta}$, with $C_{\alpha}$ being lower in modern Tianjin than Proto-Tianjin, and the complexity of the rest of the system being equal. $C_{default}$ is the base complexity of the system itself excluding individual features, including operators such as *trans* and *shift* (in Tianjin; in Songjiang *shift* resides inside the [contour] feature), and the information content of the value of features themselves (i.e. being '+' or '-').

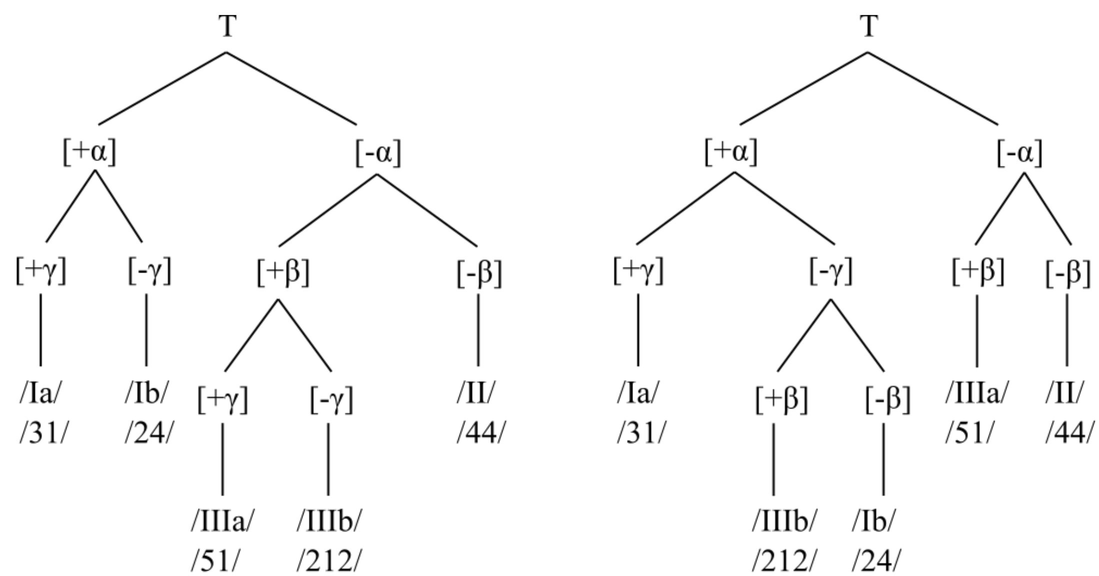
We do not include a measure of structural complexity of arranging features into a hierarchical structure, or need to directly compare the cost of a four-tone, two-feature symmetric system vs., for example, a four-tone, three-feature asymmetric system outside of directly comparing the complexity of operators. We take the view that trees are simply abstractions of the derivation of categories and of phonetic form via the application of different clusterings of operators at each stage, and thus do not incur any cost.[87]

Although Nantong is a five-tone system, the most recent reanalysis, involving a restructuring of the tree with respect to Middle Chinese tone IIIb, is in fact quite similar to that of Tianjin. As discussed in §3.2.6, since bifurcating from Middle Chinese, Nantong underwent a

[87] As discussed above, the values of the respective features themselves add $C_n$ to the total complexity, where *n* is the total number of nodes; however, this is irrelevant for the purposes of reanalysis of this type as the number of nodes is constant. Interestingly, this implies that the implementation in theory of privative features, which would not possess this (relatively small) contribution to total complexity may not necessarily be more parsimonious than binary ones. If a given model utilizing privative features requires the postulation of more features in order to account for a given inventory than a model utilizing binary features, then the extra contribution to total complexity resulting from the additional operators in a privative model may exceed the complexity contribution from feature values in a binary one.

partial merge, a tone III split, and a tone IIIb reanalysis. I focus here only on the most recent phonological change, from Proto-Nantong II to Modern Nantong, where the modern correlate of Middle Chinese tone IIIb was reanalyzed from [-α] to [+α], as complexity is only directly relevant to this change:

(117) Proto-Nantong II Modern Nantong




In both Proto- and Modern Nantong, [β] and [γ] appear to be identical in terms of phonetic content: [+β] tones are falling whereas [-β] tones not falling, and [+γ] tones are also falling whereas [-γ] tones are rising.[88] In terms of operators, by default tones are flat and the set of undifferentiated tones are, minimally, only the outputs of the *trans* operator. [β] tones contain *shift*, here a decrement operator with positive feature values triggering decrementation across the syllable. [γ] tones also contain a modified *shift* operator, *shift2*, with positive feature values triggering

88 Thus the concave tone, /212/ is both falling and rising.

incrementation and negative values triggering decrementation. Unlike Tianjin, there is both a concave tone and a rising tone, thus in order to contrastively differentiate the two we need to argue either that the concave tone, /212/ is underlyingly rising, but is contrastively low whereas the simple rising tone, /24/ is neither high nor low; or that /212/ is contrastively concave. I assume the latter position here, and account for the contour by postulating that the individual increments in *shift2* are initially smaller than in *shift*, but that the output is fed into itself recursively, incrementing the increment across the syllable, as below, in the simulated output of a syllable with an arbitrarily chosen five temporal division, or ‘ticks’, and increment values of 2 and 1 for *shift* and *shift2* respectively:[89]

| | | | Time index | | | | |
|---|---|---|---|---|---|---|---|
| | Operators | | 1 | 2 | 3 | 4 | 5 |
| a. | | (primitive) | [n | | | | ] |
| b. | *trans(*n*)* | (undifferentiated) | [n | n | n | n | n]$_\sigma$ |
| c. | *shift( trans(*n*) )* | (falling) | [n | n-2 | n-4 | n-6 | n-8]$_\sigma$ |
| d. | *shift2( trans(*n*) )* | (rising) | [n | n+1 | n+3 | n+6 | n+10]$_\sigma$ |
| e. | *shift2( shift( trans(*n*)))* | (concave) | [n | n-1 | n-1 | n | n+2]$_\sigma$ |
| f. | *inst(shift2( shift( trans(*n*)))),* | n=2, (low-concave) | [2 | 1 | 1 | 2 | 4]$_\sigma$ |

Table 26: Generation of a hypothetical low concave tone, (/214/), using two *shift* operators.

89 As with Tianjin, while the present analysis successfully generates tonal shapes, a more granular analysis becomes possible with more robust production and perceptual data.

[β] and [γ] appear to be unchanged between Proto- and Modern Nantong and, as in Tianjin, the only feature of consequence to the reanalysis is [α], which deals directly with tone height. In Modern Nantong, [α] is a simple height distinction, containing minimally only *inst*, with [+α] tones being instantiated at a contrastively low height and [-α] tones being instantiated high. There are multiple possibilities regarding the role of [α] in Proto-Nantong, however, all of them involve more complexity than Modern Nantong. The most likely possibility is that, like Tianjin, [α] in Proto-Nantong is simply a more complex height feature, with [+α] tones being central and [-α] tones peripheral, with the height of the peripheral tones determined by whether a tone is rising or not (rising [-α] tones are extra low, non-rising [-α] tones are extra high, which predicts that to speakers, /212/ in addition to being contrastively concave, is also perceptually lower than /24/). This involves the division of the pitch range into three registers, rather than two, as well as the inclusion of some sort of *check* operator, and is thus more complex.

In both Tianjin and Nantong, reanalysis is motivated by the existence of an alternate possible hierarchy with a lower complexity than was previously the case. This motivates the following:

(118) **Minimization of Complexity:** *If two or more possible hierarchies are permitted by activity patterns, choose the hierarchy with the lowest total complexity.*

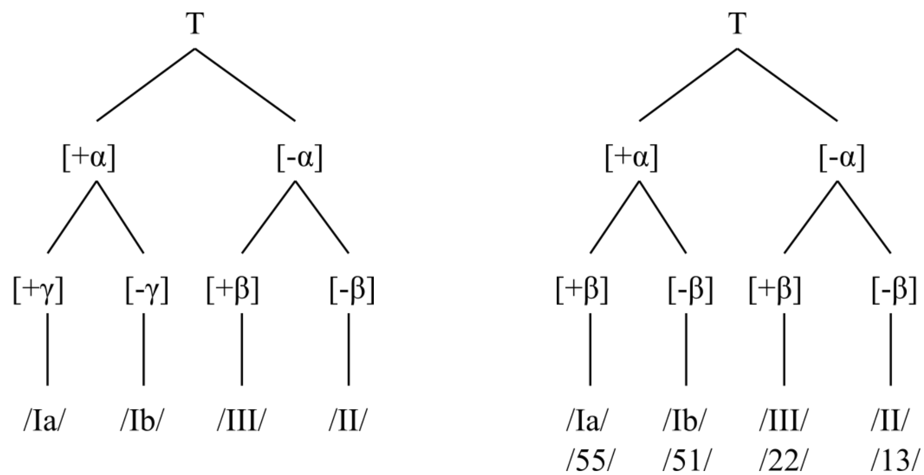
Assuming **minimization of complexity**, reanalysis in cases like Tianjin and Nantong is the consequence of learners acquiring a language that has evolved to a point where two or more distinct structural configurations are permitted by activity. When reanalysis occurs, it is simply that the new system represents an easier, or more readily postulated hypothesis for a grammar, given data available to learners, than the older one. This is a tidy solution to the problem of induction in cases like this. While there are an infinite number of potential grammars compatible with any given finite set of data, not all hypotheses are equal, and complexity metrics not only permit a solution to this issue while accounting for empirical data, they do so in a way that derives from general cognitive principles.

## 5.5 Complexity in the absence of activity: The Yue languages

The subject of this section, the Yue languages, which generally lack robust activity, provides an interesting contrast to the preceding discussion on reanalysis in the Mandarin languages, which are generally characterized by an abundance of sandhi processes. In the preceding discussion in §4.2.3, two important observations were made in a comparison of the tendencies of the tonal phonologies of Mandarin vs. Yue languages. Firstly, it was observed that the tonal inventories of most Yue languages exhibited relatively simpler tonal shapes, when compared to the Mandarin languages. Although Yue languages tend to have a larger number of tones on average, the tones themselves tend to be simpler, with more flat tones, and fewer, if any complex contoured tones (i.e. concave and convex tones). Secondly, it was observed that, relative to the Mandarin languages, the Yue languages tended to assemble their inventories into phonetically transparent hierarchical configurations. In light of the discussion in §5.2, we can now formulate these two observations in slightly more precise terms: features in the activity-poor Yue languages tend to be, on average, of lower complexity than those in the activity-rich Mandarin languages. In addition, therefore, to providing a complexity-based account for the diachronic pathways between Middle Chinese and the various modern Yue languages outlined in §4.3, an important overarching goal of this section is to outline the often direct relationship between phonological activity and feature complexity in these languages.

Zhongshan is a simple case and a suitable starting point for a discussion of complexity in the Yue languages. It is reproduced here in (119) along with a Middle Chinese hierarchy:

(119) Middle Chinese Zhongshan, [high] > [contour][90]




Zhongshan is structurally conservative, having undergone no splits or mergers, although a simpler type of reanalysis can be hypothesized: the tones in the domain of [+α] have been reanalyzed to having a [β] contrast, rather than a [γ] contrast. As discussed in §4.3.2, in §3.2.7, and elsewhere, we know that Middle Chinese tone I bifurcated into Ia and Ib on the basis of height, whereas the

---

90 Some might prefer [flat] as a label over [contour] here. Since we conceptualize features as analogous to functions which take feature values as inputs, the decision to label [α] as [flat] or [contour] is purely notational. This particular labelling is chosen here in keeping both with the convention in the rest of the paper of relegating, when possible, the modern correlates of MC TIa and TIb as contrastive sisters in the domain of [+α], and the convention of attempting to keep novel feature labels at a minimum.

contrast between tones II and III is much older, and likely contour-based. The present inventory of Zhongshan suggests, however, a single feature which performs the same work in differentiating both tone Ia from Ib and tone II from III. [β] in modern Zhongshan is thus a [contour] feature, with directionality determined by a tone height contrast: [high] tones are falling whereas [low] tones are rising (as in Barrie 2007; Yip 2001).

This is precisely what we would expect to occur over time in a system without activity, assuming a principle of **minimization of complexity**. All things being equal, it requires more information to encode three distinct features than it does to encode two. Imagine first an inventory comprised of the following four tones:

/22/ /33/ /31/ /44/

There is, strictly speaking, an infinite number of combinations of functions which generate the above four tones. However, there does not seem to be any way to encode the above inventory with two features which use less information than with the three below (there are multiple ways of organizing these three features into an hierarchy, this is one possibility):

| | /22/ | /33/ | /31/ | /44/ |
|---|---|---|---|---|
| [high] | - | - | - | + |
| [low] | + | | | |
| [falling] | | - | + | |

Table 27: Minimal complexity of a four-tone system with three features.

We thus encode both [high] and [low] with *inst* operators, with the default (unmarked) height of tonemes set to 3. [falling] is a simple *shift* operator, decrementing the tone if positive.

Although the three features postulated for this example do a fairly good job of describing this system, this particular inventory is not the simplest possible inventory that exists. Keeping in mind that all tones implicitly exhibit some variation in phonetic shape, let us imagine then that one of our tones, say /33/, which is [-high -falling] and unspecified for [low], exhibits a comparatively wider phonetic variation than other tones, and the median phonetic shape of this tone is a slightly falling tone of a level somewhat between /33/ and /44/. A learner might thus acquire something akin to [43] in their input, and is left with a task of sorting out the hierarchy—is [43] contrastively falling, and at the same height as the flat /44/, or is it contrastively flat and at the same height as the falling /31/?[91] If the latter, we end up with the hierarchy above; if the former, then the learner hypothesizes the following feature representation:

| | /22/ | /43/ | /31/ | /44/ |
|---|---|---|---|---|
| [high] | - | + | - | + |
| [falling] | - | + | + | - |

Table 28: Minimal complexity of a four-tone system with two features.

This in turn entails that the learner then interprets the height of /22/ and /31/ to have no contrastive difference, increasing the likelihood that instances in their input where /22/ and /31/ were uttered

91 We exclude the possibility that the learner interprets the tone to be either both contrastively flat and high or both contrastively falling and lower as that would entail a structural merge.

at similar heights are more likely to be interpreted as convergent with their grammar, resulting in pressure for /22/ to drift higher and/or /31/ to drift lower. This dynamic in mind, we hypothesize the following principle:

(120) ***Corollary to the minimization of complexity principle:*** *All else being equal, in the absence of external impinging factors, more complex hierarchies will evolve into less complex hierarchies.*

If we take a principle of **minimization of complexity** seriously, we would predict that, given a degree of variation in phonetic values of tones, more complex hierarchies will, given enough time, tend to evolve into less complex hierarchies in the absence of activity, tonal splits, and other external factors.

Lindblom et al. (2011) consider various perspectives on what constitutes the basis for the evolutionary, developmental and on-line adult forms of languages, citing that of Studdert-Kennedy (1998) who views linguistic forms as "emerging from prior constraints on perception, articulation and learning, according to general biological principles of self-organization". They argue that linguistic forms are constantly shaped by selectional pressures such as perceptual contrast and learnability arising from use by speakers and argue that constraints such as these likely contribute to "the formation of formal structure and the intrinsic content of sound patterns" (2011:92). It can be argued that the view of cognitive complexity advanced here represents a quantifiable metric which directly determines learnability, accounts for a wide variety of feature economy effects, and is already motivated by a host of evidence from other cognitive domains. The minimization of complexity, for example, accounts both for why, as Maddieson (1984:16) asked, / i ẽ a̤ o̰ uˤ/ is not typologically favoured over /i e a o u/ if maximization of perceptual contrast is the prime factor determining inventory membership; and why, as Ohala (1979) points out, consonant inventories seem to make maximal use of available distinctive features rather than exhibit maximal perceptual difference, instead (see also Lindblom et al. 2011:93).

Similarly, minimization of complexity and maximization of transfer can help account for and further contextualize what Clements (2003) describes as the tendency of phonologies towards feature economy: the maximization of speech sounds over features. According to Clements, "languages tend to maximize the combinatorial possibilities of features across the inventory of speech sounds: features used once in a system tend to be used again" (2003:287). Clements states that the notion of economy first appeared in work by de Groot (1931), who argues that "there is a tendency to employ certain accompanying phoneme properties more than once; one can speak of a tendency towards economy" (1931:121). This notion was built upon by Martinet who made the argument that "a feature-based phoneme system is more economical than one in which each phoneme involves an entirely unique articulation (see also Clements 2003:293): "further economy is achieved by making phonemes result from combinations of non-successive phonic features, which further reduce the number of basic elements" (Martinet 1955:95). In somewhat striking accordance with results here, Clements (2003) finds in a large survey of 451 languages that feature economy works only on phonologically *active* (2003:328) features, and that the tendency of sound systems to evolve towards more simple representations likely operates at a higher cognitive level, involving "the same principles of category formation and generalization that are at work in other areas of grammar" (2003:329).

The corollary to the minimization of complexity principle also accounts for the reanalysis in Zhongshan: upward drift in the modern correlate of Ib allowed what was once a height distinction in the domain of [+α] to be reanalyzed as the same contour distinction that exists in the domain of [-α], resulting in a less complex system, and satisfying both the **minimization of complexity** and **maximization of transfer** principles. This also implies that, once a tonal system reaches a local complexity minimum, any drift results in a more complex mapping of tones, and is thus more likely to be rejected by learners as atypical utterances, derived from the underlying, less complex form. Therefore, although an inventory such as {/55/, /51/, /22/, /21/} requires less information than Zhongshan's current one (as we would not need the equivalent of a *check* operator which instantiates contour direction based on whether a tone is high or low), there is no feasible way for tonal drift to yield two falling tones or two rising tones without first traversing through a more complex inventory. Zhongshan's current contrastive arrangement is thus more or less stable, yielding the empirical prediction that any further reanalyses must be motivated by strictly external factors.

Notions of simplicity or complexity are inescapable issues that any theory of phonology must contend with. Early generative grammar, following the Prague school, utilized features as a means to relate the specifics of inventories to what was assumed to be a universal phonetic vocabulary (Chomsky and Halle 1968. See also Battistella 1995:63). Features were also incorporated into an evaluation metric, conceptualized as a procedure which determines which of the grammars, consistent with linguistic data, is to be selected as the grammar of the language (Chomsky and Halle 1968:30). Incorporating an evaluation metric into Universal Grammar presents a candidate solution to the problem of induction in acquisition, by constraining the otherwise mathematically infinite set of possible grammars which generate linguistic data, selecting the most highly valued or 'simplest' one. In SPE, the evaluation metric stipulated that the value of a candidate grammar was the inverse of the number of features required by the grammar.

It turns out, however, that this initial formulation has the side effect of rendering implausible, less typologically common rules and segment inventories as highly valued as natural and frequent ones. To contend with this issue, a set of 39 supposedly universal marking constraints were introduced, such that the unmarked value of a feature did not contribute to the value of a grammar for evaluation purposes (Chomsky and Halle 1968: 402-408). These rules were construed as stipulated in UG, being "universal…conventions for the interpretation of a grammar" (1968:403). This in turn, however, led to additional difficulties, as Battistella (1995:64) writes: "summing the values of features could not directly account for the fact that some features are more likely than others or for the fact that phonological systems have a certain minimal complexity and symmetry", and the style of markedness advanced in SPE subsequently did not receive much attention.

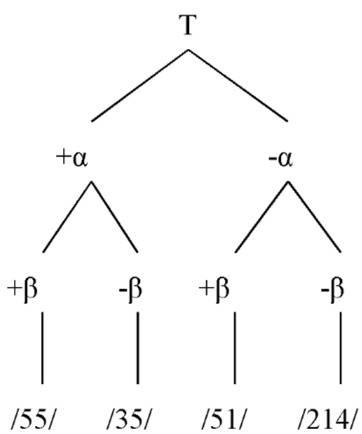
In the preceding discussion, I made a point of underscoring the idea that activity can have a profound effect on the tendency of more complex inventories to evolve into less complex ones. To illustrate how this is the case, consider the Beijing Mandarin inventory:

| | /55/ | /35/ | /51/ | /214/ |
|---|---|---|---|---|
| [α] | + | + | - | - |
| [β] | + | - | + | - |

Table 29: Beijing Mandarin inventory

As discussed in §3.2.1, Beijing possesses two robust sandhi processes: /214/ → 35 / __ /214/ and the optional /35/ → 55 / {/35/,/55/} __ T, and thus the only two-feature hierarchy possible (assuming abstract features, no specific ordering) is reproduced below:

(121) Beijing Mandarin




Feature labelling necessarily entails more complexity than Zhongshan. In §3.2.3 I speculated that we might label [α] as a falling/not falling contrast, with [-α] tones all exhibiting a falling element, and [β] as a height contrast with [+β] tones being contrastively high, noting that this required the stipulation that all low tones are rising by default. Implementing this with operators requires a fair amount of extra complexity. The [α] feature therefore encodes a simple *shift* operator, triggering decrementation in tones with a negative value. The [β] feature is where we encode both height and rising contour, requiring an *inst* operator for height, and a modified *shift2* operator, as in Nantong, which generates simple rising tone, /35/ in the absence of other contrastive contours, and a concave tones, /214/ in tones which have already been subjected to decrementation across the syllable. Again, as with Nantong, this is accomplished by our operator shifting the height of the tone upwards in progressively larger increments towards the right syllable edge, yielding an initially downward sloping tone when the static decrementation of *shift* is greater than the increasing incrementation of *shift2*, progressing to an upward sloping tone when the incrementation of *shift2* exceeds the decrementation of *shift.*

At any rate, sandhi processes in Beijing lock the inventory into the only possible two-feature hierarchy above (again, assuming abstract features and no specific ordering). Were Beijing not locked into the present configuration, there are multiple ways for the drift of single tones to reduce complexity. If, for example, /214/ were to lose its contour, becoming contrastively flat and low, we would expect a reanalysis to the following, much simpler hierarchy, resembling Zhongshan:

| | /55/ | /35/ | /51/ | /22/ |
|---|---|---|---|---|
| [high] | + | - | + | - |
| [contour] | - | + | + | - |

Table 30: Beijing Mandarin simplified tonal inventory, /214/ ⟶ /22/

Likewise, the loss of contour in /35/ accomplishes a similar effect, however, this is only possible if the loss of contour in /35/ triggers a second reanalysis, of /214/ to a simple rising contour, /24/, as below:

| | /55/ | /33/ | /51/ | /24/ |
|---|---|---|---|---|
| [high] | + | - | + | - |
| [contour] | - | - | + | + |

Table 31: Beijing Mandarin simplified tonal inventory, /35/ ⟶ /33/, /214/ ⟶ /24/

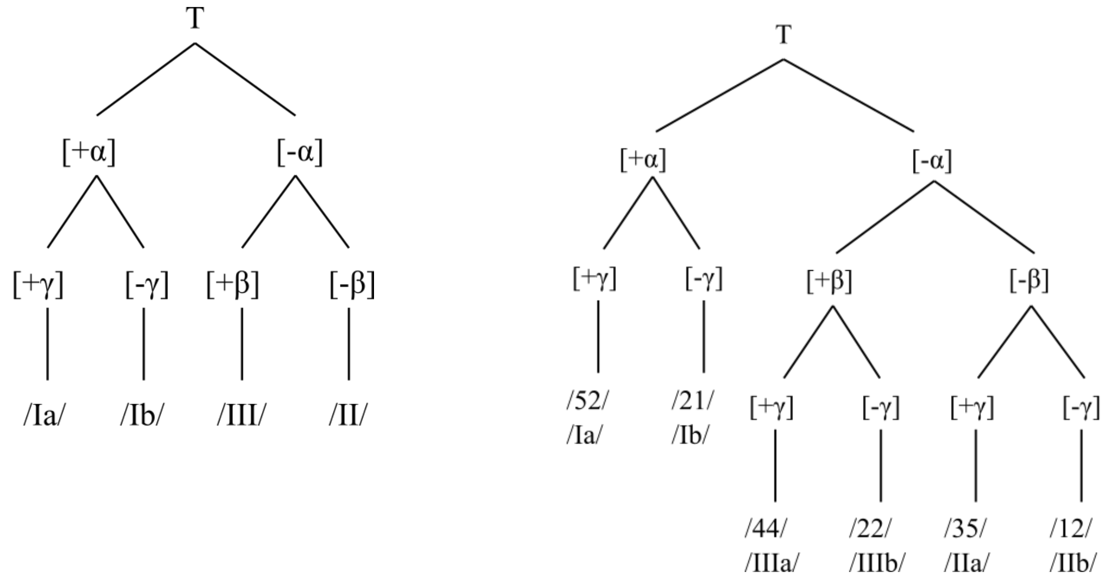
With Beijing's inventory of sandhi processes however, only the second of these two is possible: The sandhi process linking /55/ and /35/ and the process linking /22/ (formerly /214/) and

/35/ both entail the change of a single feature, and there is no configuration which simultaneously allows both processes and permits a [+/- high] and [+/- contour] set of contrasts if tonal drift permits the reanalysis of /214/ as /22/.

Although the precise mechanisms underlying phonetic drift are not well understood, the evidence thus far presented here suggests that phonological activity constrains the possible avenues by which an inventory can be reanalyzed. Phonological activity provides to learners an explicit structural link between two inventory members. This explicitness directly influences the task of generating a phonology. In doing so it also serves to restrict the possible set of remaining unspecified relations. In languages with a high degree of activity, this may come at the cost of being unable to encode the most parsimonious set of phonetic relations, with both synchronic and diachronic consequences. In addition, were some phonology-external force to put pressure to drift on a tone in an inventory that is already gridded in by activity to a single possible configuration, it is more likely than not that the encoding simply becomes more complex in response: as the hierarchy is already explicitly determined by activity, there is only weak, if any compensatory pressure for other tones to drift in response as phonetic shape is not the primary determinant of phonological relations. Put another way, the lack of activity in languages both permits and necessitates a more dynamic restructuring of contrastive relations that is in sync with phonetic change, and we would expect a high degree of compensatory, synchronized drift across the inventory in response to pressure on a single tone. It is suggested that this is the primary reason that the activity-poor Yue languages tend to, as a class, exhibit comparatively simple tonal shapes and inventories which are remarkably transparent with respect to historic categories.

Cenxi is also a very simple case which exemplifies some of the above principles. It is reproduced below, with category representation alongside a Middle Chinese hierarchy:

(122) Middle Chinese — Cenxi, [falling] > [flat] > [high]

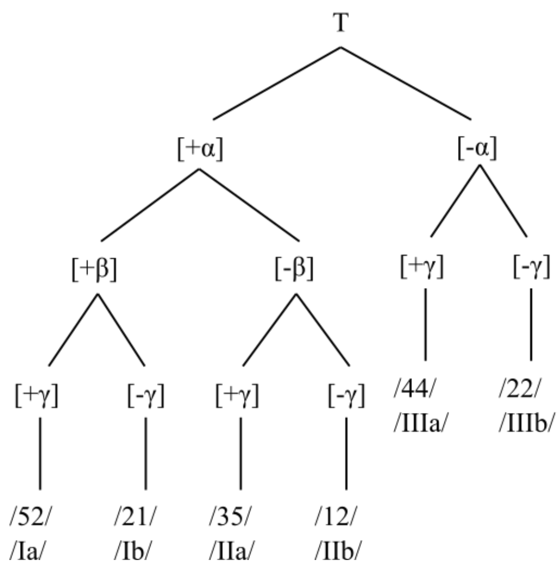



As discussed above, assuming the hierarchy in (122), Cenxi is a simple expansion of the Middle Chinese hierarchy, with tones II and III undergoing a split. [+α] tones are [falling] whereas [-α] tones are [-falling]. [β] is similarly a contour contrast, with [+β] tones being flat and [-β] tones being rising. [γ] is a height contrast and [+γ] tones are [high] and [-γ] tones are [low].

Feature encoding is straightforward in this scenario. [γ] consists of a simple *inst* operator, with the pitch range being divided into high and low, and the height of the resultant tone being determined by the value of the feature. [β] consists of a *shift* operator, with positive values triggering no shift, and negative values triggering incrementation. The [α] feature similarly is composed of a decrementation operator, *shift2*. It is worth noting that both *shift* and *shift2* in this conceptualization are static increment operators[92] and of equivalent complexity to each other.

There is a second, equally parsimonious possibility. Rather than having two distinct *shift* operators, tones in Cenxi may first be divided by contour, then by directionality, and finally by height, as below:

(123) Cenxi, [contour] > [falling] > [high]




This hierarchy encodes features slightly differently. Here, [γ] ([high]) also consists of a simple *inst* operator, and [α] ([contour]), similarly consists of a *shift* operator. However, the encoding of this *shift* operator may be slightly more complex—it is undifferentiated in terms of directionality, rather than being static, and thus requires an intermediate abstract contour for an extra step in the derivation. [β] ([falling]) however is less complex in this formulation: it is a simple

92 That is, both *shift* and *shift2* simply increment the pitch across the syllable at a constant rate. This is unlike *shift2* in Nantong, which is more complex, incrementing in progressively larger steps towards the right boundary.

binary switch, *inv*, determining directionality, with positive values triggering incrementation and negative values decrementation.

It is worth noting that, since Cenxi has no activity, every single hierarchical configuration of tones is at least notionally possible. Structural reanalyses, where only the contrastive relationships between tones, and not the phonetic shape of the tones themselves change is, in phonologies without activity, simply an automatic consequence of tonal drift reaching some critical inflection point. A particular encoding of a feature, or a particular arrangement of features into a system of oppositions and similarities is a hypothesis, the direct result of the cognitive apparatus compressing information for later reconstruction. Hierarchies, being individually constructed by learners, aren't directly passed between speakers, and reanalysis is not something which occurs to a hierarchy directly. Rather, reanalysis is conceptualizable as a change to which hypothesis is the least complex algorithm that generates the inventory.

Comparing the two hierarchies above therefore, *inst* is of equivalent complexity in both (122) and (123). The hierarchy in (123) saves some complexity in the encoding of an *inv* operator, which is a simple binary switch, rather than a second *shift* operator, which is more complex, however this comes at the tradeoff of having a likely more complex single *shift* operator. Since we do not have access to the specifics of the encoding of the cognitive apparatus, we are not able to decide whether a more complex *shift* plus an *inv* operator is more or less complex than the encoding of two simpler static *shift* operators. If it is more complex, we predict the hierarchy in (122), if less, we predict (123).[93] Cenxi is also at a local complexity minimum, which we can now define as follows:

(124) ***Complexity Minimum:*** *A phonological subsystem is at a complexity minimum if there is no possibility for the simple drift of a single inventory member to yield a less complex hierarchy.*

As with Zhongshan, there is no drift in Cenxi which yields a less complex inventory or hierarchy without massive restructuring. If we take seriously the observation (Maddieson 1978; Yip 1980; Bao 1990, 1999) that tonal languages universally do not exhibit more than four level tones, the only simpler hierarchy in a 6-tone language requires four flat and two falling tones (e.g. [height] > [falling] > [height], with two binary *inst* operators and one simple static *shift* operator), which is not a configuration that is easily accessible to Cenxi through simple drift.[94]

Binyang, similarly, is a good case of a language which exemplifies the principles discussed above, and is a good illustration, alongside Cenxi and Zhongshan, of how languages minimize complexity while preserving contrastive relations via synchronization of tonal drift across an inventory.

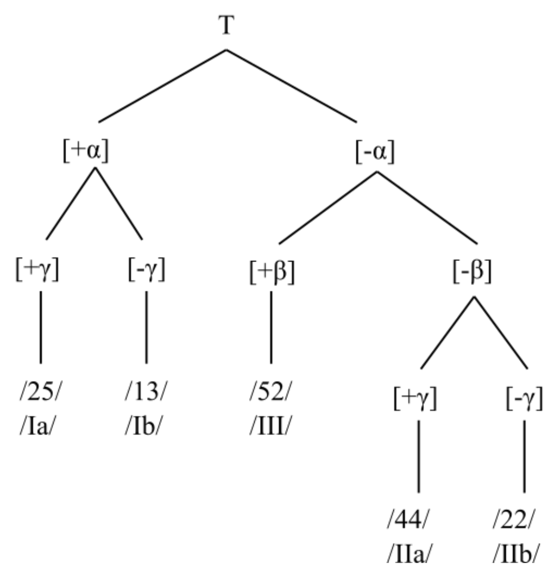
---

93 It is altogether possible that the two co-occur in a given population as variation in cases like these likely has a dramatic effect when an inventory is close to a critical inflection point where multiple algorithms of similar complexity can generate it.

94 For this to occur, a highly implausible set of circumstance would need to occur simultaneously. Zhang (2002) observes that a language must have falling tones if it has rising tones, thus if a language only has two contoured tones, the simplest possible configuration is two falling tones, since a falling and a rising tone requires a more complex set of operators. This in mind, a two-height one-contour hierarchy for Cenxi would require the simultaneous drift of: /52/ → /42/, /21/ → /31/, /35/ → /33/, /12/ → /11/.

To start with, the hierarchy for Binyang, which is at a local complexity minimum, is reproduced below. [α] is a [rising] feature, encoded with a static *shift* operator, [β] is a [falling] feature, also encoded with a second *shift* operator, and [γ] is a height feature, encoded with *inst.*[95]

(125) Binyang

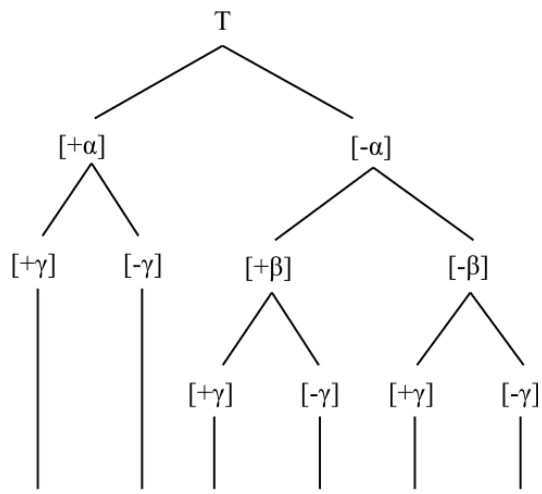



As can be seen, there is no evidence of reanalysis, and the familiar TIa/TIb vs. TIII/TII configuration originating from Middle Chinese is present in the Modern Binyang hierarchy. Binyang exhibits no activity, thus the only mechanism via which structural relations are preserved is phonetic shape. It was suggested above that in languages without activity, compensatory or synchronized drift occurs to maintain structural relations. This is made clear in the composite

95 As with Cenxi, also possible is a hierarchy with a more complex contour distinction, *shift*, which generates an undifferentiated contour with directionality determined by a subordinate binary switch *inv*, and an *inst* operator determining height. It is not decidable which of these trees is more complex, so we postulate the more conservative hierarchy in (125).

hierarchy below. All three languages are descended from Late Middle Chinese / Proto-Yue and have the same modern structural skeleton, despite the presence of significant drift.

(126) Zhongshan, Cenxi, Binyang with Middle Chinese categories[96]



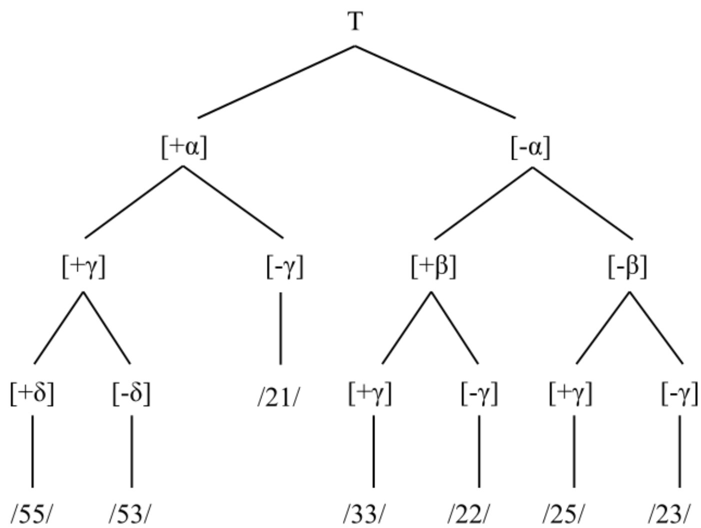

| Category | /Ia/ | /Ib/ | /IIIa/ | /IIIb/ | /IIa/ | /IIb/ |
|---|---|---|---|---|---|---|
| Zhongshan | /55/ | /51/ | /22/ | /22/ | /13/ | /13/ |
| Cenxi | /52/ | /21/ | /44/ | /22/ | /35/ | /12/ |
| Binyang | /25/ | /13/ | /52/ | /52/ | /44/ | /22/ |

We would not expect patterning such as this to happen by chance without a principle of complexity minimization. The hierarchy with the lowest complexity in each language exhibits similar structural relations in each other language as well as in Middle Chinese. At the same time, structural relations are preserved despite phonetic drift yielding drastically different tonal shapes in each language. The only way to account for this is to postulate the synchronization of drift motivated by the least complex formulation of contrastive relations at each point in time.

96 As discussed above, Zhongshan possesses a [β], rather than a [γ] feature in the domain of [+α].

Hong Kong and Guangzhou Cantonese also demonstrate complexity minimization.

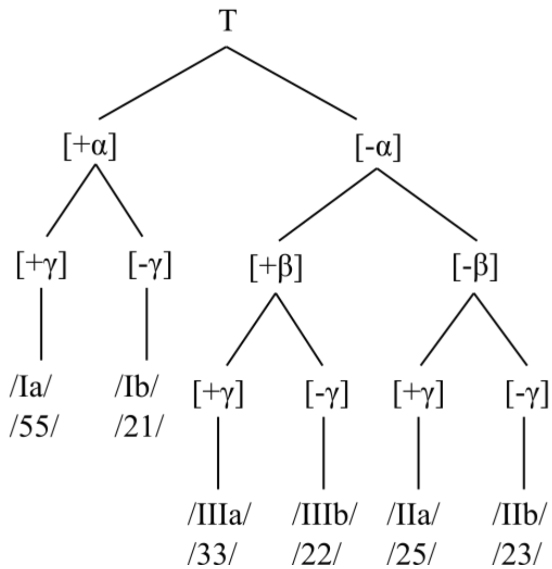
(127) Guangzhou Cantonese




As discussed in §4.2.3, there is significant evidence for supposing the structure above, including: i) partial mergers between /33/ and /22/ (Ou 2012); ii) a complete merger between /55/ and /53/ in most younger speakers (Barrie 2007; Matthews and Yip 1994); iii) the existence of a sandhi process linking /55/ and /53/ in all speakers which distinguish the two tones; iv) diachronic evidence from Middle Chinese and other Yue languages which share similar interrelations between categories.[97] This basic structure also serves as a protoform for Modern Hong Kong Cantonese, where /55/ and /53/ have merged completely:

97 There is no three-feature system which accounts for the presence of both mergers, the sandhi process, and phonetic shape.

(128) Hong Kong Cantonese I, [peripheral] > [rising] > [high]




As with Guangzhou, there is also evidence in Hong Kong Cantonese for a partial merger between /25/ and /23/ (Bauer et al. 2003; Kei et al. 2002; So 1996; So and Varley 1991), thus, at a minimum, we would expect /25/ and /23/ to be contrastive sisters.

If it is indeed the case that the four remaining tones, /55/, /33/, /22/, and /21/ can be freely assigned contrastive relations by learners based purely upon phonetic form, we expect an additional reanalysis based on the principle of complexity minimization. Assuming the structure in (128), [γ] is a height feature, straightforwardly implemented as a binary switch with *inst*, [β] is a contour feature, implemented with *shift*, and [α] is a peripheral feature. The implementation of a central/peripheral contrast in making one cut of a four-way height distinction, like a simple height feature, requires a binary switch, *inst*. However the combination of a central/peripheral contrast with a height contrast introduces slightly more complexity than the combination of two height features. The reason lies in the way that the two features are combined to generate height. In the case of two height features, say [+/- high] and [+/- upper], the simple merge of two bits gives the desired result. In the case of a height feature and a peripheral feature, an additional operation is required. To show how this is the case, first we represent feature values as single binary digits, with positive values represented as 1, and negative values as 0. Then we merge the two binary numbers to compose a two digit string, as below:

(129) [upper] > [high], value of binary flags for two *inst* operators

| | **merged binary string** | **value** | **tone height (binary)** |
|---|---|---|---|
| [+upper] [+high] | 1 1 | 3 | 3 (1 1) |
| [+upper] [-high] | 1 0 | 2 | 2 (1 0) |
| [-upper] [+high] | 0 1 | 1 | 1 (0 1) |
| [-upper] [-high] | 0 0 | 0 | 0 (0 0) |

With the [upper] > [high] hierarchy, there is an isomorphic relationship between the merged value of the two features, and the concomitant tone height. The features are of minimum complexity, as they are essentially empty save for the value of the features themselves. This is in accord with the observation that height will be the contrastive split in any language with only two contrastive tones (see Maddieson 1978; Yip 1980; Bao 1990, 1999; and Zhang 2002 for a discussion of tonal universals). Utilizing a peripheral contrast however complicates things slightly:

(130) [peripheral] > [high], value of binary flags for two *inst* operators

| | **merged binary string** | **value** | **tone height (binary)** |
|---|---|---|---|
| [+peripheral] [+high] | 1 1 | 3 | 3 (1 1) |
| [+peripheral] [-high] | 1 0 | 2 | 0 (0 0) |
| [-peripheral] [+high] | 0 1 | 1 | 2 (1 0) |
| [-peripheral] [-high] | 0 0 | 0 | 1 (0 1) |

With a [peripheral] feature, there is no longer an isomorphic relationship between the merged feature values and the binary representation of tone height, so a second operation needs to be performed on the string.[98] [99] The *information distance* (Li and Vitanyi 1992; Bennet et al. 1998), between two objects, *x* and *y*, can be defined as the length of the shortest algorithm which converts *x* into *y*. In an [upper] > [high] system, the information distance between feature values and tone height is zero, but is greater than zero in a [peripheral] > [high] system, rendering it a more complex system overall. Interestingly, this also accords with an identical tendency reported for vowel systems. Rice (2007) provides evidence that central vowels generally pattern as unmarked, for example, serving as a target, but never a trigger for assimilation, following the

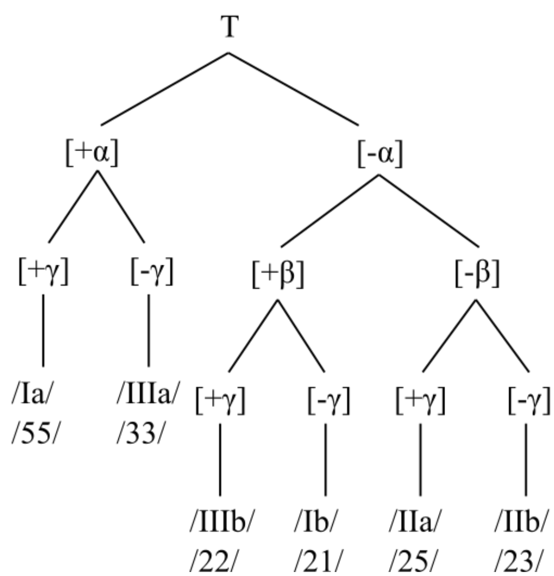
98 The least complex operation here is likely the cognitive equivalent of the simple application of a NOT gate to the second digit of our binary string, if the first digit is 0, and copying it if the first digit is 1, with the first digit of the output being the second digit of the input, and the second digit of the output being the output of either the NOT or copy operation.

99 Of course, there is no evidence that the cognitive apparatus encodes tone height in binary prior to outputting it phonetically. However, these examples illustrate that a [peripheral] contrast represents a formally distinct encoding system from a [high] contrast, and the combination of the two systems requires the introduction of additional complexity. It is impossible to encode height solely with a [peripheral] contrast.

assumption that there is no [central] feature (see also Dresher 2009).[100] This is what we would expect in a system attempting to minimize complexity for exactly the same reasons outlined above for tone.

This has implications for the modern form of Hong Kong Cantonese. Since we expect learners to construct the least complex hierarchy, we thus predict a further restructuring of the non-rising subsystem as follows:

(131) Hong Kong Cantonese, restructured, [upper] > [rising] > [high]




With this revised hierarchy, [α] is an [upper] feature, [β] is a [rising] feature, and [γ] is a [high] feature and complexity is at a minimum.

---

[100] While central/peripheral distinctions clearly do exist in tonal systems, the evidence presented suggests they are dispreferred. Given the monodimensional nature of tone, tonal systems cannot make recourse to other acoustic dimensions to reduce complexity, likely resulting in higher total complexity of individual features.

It is interesting that /21/, which we presume to be contrastively flat, is nevertheless phonetically falling, and formerly a contrastive sister to the TIa category, which first split into a falling and flat variant, later merging again. This suggests a broader pattern of a gradual loss of contrastively falling contours overall (recalling that Proto-Hong Kong/Guangzhou possessed two contrastive tones, now merged, in the modern TIa category, as discussed in §4.2.3). If so, this implies that /21/ was likely formerly mid height and contrastively falling, and the [α] contrast in (128) was formerly simply a register contrast, with [+α] tones being [upper] and falling if contoured, and [-α] tones being lower, and rising if contoured (as in Barrie 2007; Yip 2001). This in turn implies that there never was a higher complexity [peripheral] feature in Hong Kong Cantonese to begin with.[101]

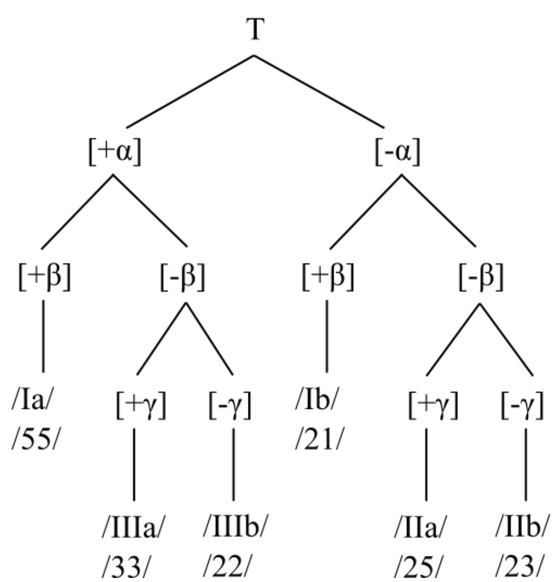
Up until now, relatively little has been said regarding the contrastive status of tone Ib, /21/ in Hong Kong Cantonese. While our analysis above generally treats it as flat, there is significant disagreement in the literature regarding whether to treat it as contrastively low and falling (Bauer and Benedict 1997; Mok and Wong 2010; Lee 2012; Mok et al. 2013), or as extra-low and flat (Barrie 2007; Ou 2012; Matthews and Yip 1994; Yip 2002;  Tse 1992). Phonetic measurements of the tone itself by individual speakers (Mok and Wong 2010; Mok et al. 2013) show a clear falling contour; however, other authors (Matthews and Yip 2011) report both that there is significant variation in the production of the tone, which is “sometimes pronounced at an even pitch: 11”

---

101 In a similar vein, we would also expect a reanalysis in Standard and Doushan Taishan such that a peripheral contrast is replaced by a simpler height contrast (see Appendix 3). Unlike Hong Kong Cantonese, the significantly more complex diachronic pathway that Taishan underwent renders undecidable the question of how and in which categories tonal drift conditioned the final Taishan restructuring. More data is needed to resolve this question.

(2011: 28-29), and also that variation seems to pattern at least with gender (2011:29). Assuming the presence of some sociolinguistically relevant variation in the contour of this tone seems reasonable, given the facts, and also fits within the general pattern of a gradual loss of contrastively falling tones in Hong Kong Cantonese hypothesized above. Whether /21/ is treated as contrastively falling or flat by learners has profound consequences for how the hierarchy is constructed. If /21/ is treated as flat, we expect the hierarchy above. If it is treated as contrastively falling, however, /21/ needs to pattern with the other contoured tones, rather than the flat ones, and the least complex hierarchy becomes the following:

(132) Hong Kong Cantonese, /21/ as falling




[α] here has been reanalyzed as a [contour] feature, implemented with an undifferentiated *shift* operator. [β] is a register feature, for example [upper], implemented with a binary flag, either *inst* or *inv*, which, for positive values sets flat tones to the upper register, and contoured tones to be falling, such that directionality of a contour may be stipulated by an already extant register contrast (as in Barrie 2007; Yip 2001). [γ] is a simple [high] feature, and is also implemented with an *inst* operator.

It is worth noting that the interpretation of /21/ as either contrastively falling or flat directly affects which tones in the hierarchy share features with which other tones. This should, in principle, affect patterns of variation and diachronic change. One such pattern is the report (Mok and Wong 2010; Mok et al. 2013) of a highly inconsistent partial merger amongst some younger speakers between /22/ and /33/. At least some of the evidence for this merger appears to be anecdotal, and there doesn't seem to be any evidence of full neutralization in any speakers. That said, we would expect this merger to only be possible if we assume the two tones differ by a single feature, which is the case in the hierarchy which treats /21/ as contrastively falling, but not the one which treats it as contrastively flat. While interesting, more research on this phenomenon is clearly needed.

## 5.6 Summary

The preceding chapter outlined a conceptual framework for identifying and formalizing complexity in tonal systems, and incorporating such into the broader theoretical framework of the Contrastive Hierarchy. I discussed a variety of ways in which the cognitive preference to minimize complexity affects diachronic change, and showed that reanalysis is the predictable consequence of tonal systems reaching an inflection point where a new phonology (either a new feature labelling or new set of structural relations between inventory members, or some combination of both) exhibits a lower complexity than an old one.

# 6

## Concluding Remarks

I have, in the preceding chapters, (1) shown that our proposed framework of contrastive features may be successfully implemented for tonal systems in 13 Chinese languages, (2) accounted for the modern structure of each of these languages by postulating a small number of contrast shifts and reanalyses from Middle Chinese, and (3) argued that phonetic naturalness plays a significant role in conditioning diachronic change, and that phonetic naturalness can be conceptualized as derivative from the general cognitive preference for low complexity representations.

In Chapter 2, I observed that arguments against tonal features generally rest on the assumption that features are drawn from a small universal set (Clements et al. 2010; Hyman 2010, 2011) or organized according to some universal, static set of conventions (e.g. universal markedness constraints in Wang (1967) and Chomsky and Halle (1968)). Permitting language-specific features resolves many of the supposed issues with these analyses, and I hypothesized that the framework associated with the contrastive hierarchy provides a promising means to analyze tonal systems.

In Chapters 3 and 4, I derived feature hierarchies for 13 Chinese languages in the Yue and Mandarin groups and then traced their development from a reconstructed Middle Chinese utilizing a small number of principles: i) that phonological activity involves only the change in value of single features; ii) that splits result in contrastive sisters, differing by a single feature at the most subordinate level of our hierarchy and that full mergers similarly only occur between two contrastive sisters; iii) that phonetic naturalness motivates structural relations in the absence of phonological evidence (e.g. activity) to the contrary; and iv) that structural reanalyses are similarly motivated via phonetic naturalness considerations. I argued that the construction of contrastive hierarchies is an organizing principle which influences and places constraints on acquisition and change.

In Chapter 5 I presented evidence that the cognitive apparatus exhibits an inherent preference for the least algorithmically complex candidate representations of data, and all evidence indicates that this preference is general, cutting across multiple cognitive domains, and is a core requisite of pattern recognition, and information compression and reconstruction (Attneave 1954; Barlow 1959; Garner 1970, 1974; Li and Vitányi 1997; Chater 1999; Falk and Konold 1997; Leyton 1984, 1986a, 1986b, 1986c, 1992; 2001; Pothos and Chater 2002; Chater and Vitányi 2003; Desalles 2006, 2010, 2012; Saillenfest and Desalles 2012; Peter van der Helm, to appear, amongst others). Theories which are able to decide between the complexity of multiple competing candidate representations within a particular domain have the desirable characteristic of possessing notionally bounded and quantifiable parsimony, and the notion of Kolmogorov complexity or algorithmic entropy provides a fertile theoretic foundation towards this end (Kolmogorov 1963, 1968; Solomonoff 1978; Li and Vitányi 1997; Chater and Vitányi 2003; Desalles 2006, 2010; Saillenfest and Desalles 2012).

I have also outlined a conceptual framework for incorporating notions of complexity into the broader theoretic framework of the Contrastive Hierarchy (Dresher 2009). I argued, building on research by Leyton (2001), that features can be conceptualized as bundles of operators, *transfer groups,* which are applied sequentially onto structure, *fiber groups,* with inventories as the eventual output, and that a contrastive hierarchy is the abstract representation of all possible derivational pathways utilized by a given inventory. I noted that the framework associated with the Contrastive Hierarchy is already compatible with a central precept of minimizing complexity, namely the **maximization of transfer** (Leyton 2001:4), which makes one part of the generative sequence a transfer of another part of the generative sequence whenever possible, via the hierarchical application of features. This, in principle, maximizes theoretical parsimony over any stipulative, non-hierarchical, or non-compositional theoretical alternative.

With respect to tonal systems, I discussed two different types of reanalysis: the reanalysis of structural relations, and the reanalysis of feature content. I presented evidence that showed that with both types, reanalysis is the predictable consequence of tonal systems reaching an inflection point where a new phonology (either a new feature labelling or new set of structural relations between inventory members, or some combination of both) exhibits a lower complexity than an old one. This warranted the postulation of the **Minimization of Complexity** principle: if two or more possible hierarchies are permitted by activity patterns, choose the hierarchy with the lowest total complexity. I argued that minimization of complexity, like the construction of contrastive hierarchies is an organizing principle which permits the formulation of empirical predictions and accounts for a number of synchronic and diachronic phenomena.

In specific, I showed that, in Tianjin and Nantong Mandarin, structural reanalyses were motivated by a complexity drop. I also demonstrated that phonological activity, which serves to restrict the structural possibilities available to learners, has the effect of increasing the complexity of features as a result, via the limitation of avenues to reduce complexity introduced by tonal drift. In a similar vein, it was shown that the phonetic inventories of Cenxi, Zhongshan, Binyang, Hong Kong Cantonese and Taishan, none of which exhibit sandhi processes, are all at local complexity minima, and that the maintenance of low complexity inventories shaped diachronic development from Middle Chinese. Guangzhou, the only Yue language in this data set which exhibits phonological activity, is also the only Yue language which is not at a complexity minimum.

The evidence suggests that complexity underlies what phonologists have come to call phonetic naturalness: a phonetically natural arrangement of features is one where complexity is minimized. This makes sense intuitively: the preference for simplicity is domain general, and the same preference which minimizes complexity in representations also allows its recognition by analysts. It also makes sense typologically, and the reason is twofold. Firstly, simpler classes are more likely because the cognitive apparatus prefers simple representations: all things being equal we assume that this preference imposes an upper limit on the total complexity of inventories and both mitigates drift away from natural arrangements due to external (for example, sociolinguistic) factors, and induces gradual motion towards natural arrangements. Secondly, all things being equal, a given class will be more common as its complexity decreases: the probability that two typologically unrelated languages share a given set of structures decreases as the complexity of those structures increases.[102] Taken together, the evidence provides an account for what Lindblom et al. (2011) describe as the tendency for languages to be continuously shaped by selectional pressures, such as learnability, and what Clements (2003) describes as the tendency towards maximization of feature economy.

A number of areas will benefit from a richer data set and additional research. In the case of Taishan, we were unable to motivate the entire sequence of structural reanalyses since Middle Chinese due to the highly complex sequence of changes. While Taishan presents no counterevidence to any of the principles advanced here, more detailed historic evidence from either Taishan itself, or closely related dialects, would likely help to enrich the diachronic picture we have of this language. With respect to Yangjiang, more robust data on the modern Yangjiang inventory is required to conduct any detailed complexity-based analysis, and at present there are simply too many plausible structural possibilities to decide on a single analysis.

Pingyao and Yantai are both interesting cases requiring more detailed investigation. Both of these languages possess inventories which contain two members which are merged in citation

---

[102] This is simple to show. The complexity of generating a single digit number in base 10 ($\log_2 10 \approx 3.32$) is half that of a two-digit number of the same base ($\log_2 100 \approx 6.64$). Similarly, the probability that two random single digit numbers correspond is an order of magnitude greater ($\approx 1/10$) than the probability that any two given random two-digit numbers correspond ($\approx 1/100$). As the complexity of a given object increases, the probability that another random object of the same complexity corresponds with it decreases.

form, but which behave differently in sandhi contexts. It is unclear whether this represents a merge in progress, or if the behavior is stable. The latter case introduces an intriguing nuance to a complexity-based analysis: since we presume that the two tones differ by a single feature, what is the precise encoding of this feature? Admittedly this is not only an issue of theoretical interest to complexity-based analyses. More robust data on these languages however is clearly needed as we do not have acoustic or perceptual data of the citation tones in question, or even sociolinguistic data regarding extant tonal variables and cannot completely rule out the possibility that consistent differences exist in citation form.

While this study has focused on tone, the results, theory, and architectural assumptions of the present work do not preclude treating segment inventories with similar analyses as deployed here. Complexity minimization and transfer maximization are both domain-general processes, and the ‘universal’ element of phonology is presumed here to comprise only the formation of abstract categories via these principles, yielding contrastive hierarchies as emergent products. Indeed, the observation noted above, that vowel systems tend to disfavor features such as [mid] or [central], as well as the general tendency for inventories to maximize feature economy, follows directly from the minimization of complexity principle and an analysis of segmental phonologies is thus a potentially fruitful area for future study.

The relationship between phonological activity and phonetic naturalness, too, seems an interesting area for future research. The means to formalize complexity advanced here presents a new analytical tool for phonologists towards quantifying phonetic naturalness, opening up avenues of inquiry which were heretofore unavailable. While the existence of some relationship between activity and naturalness is more or less consensus (Mielke 2004:229), to my knowledge, the present study is the first to postulate both a specific, general mechanism which results in the crosslinguistic tendency for inventories with more activity to be less phonetically natural, and a non-stipulative means for quantifying complexity. It would thus be interesting to see to what extent this relationship is borne out in the inventories, both tonal and segmental, of a wider set of languages.

# *References*


Anderson, Stephen R. (1978). Tone features. In Victoria A. Fromkin (ed.), *Tone: a Linguistic Survey*. 133-176. Academic Press.

Anderson, Stephen R. (1988). Morphological Change. In Frederick Newmeyer [ed.], *Linguistics: The Cambridge Survey* (vol. 1). 324-362. Cambridge: Cambridge University Press.

Ao, Benjamin Xiaoping. (1993). *Phonetics and Phonology of Nantong Chinese*. Ph.D. dissertation, Ohio State University.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychol. Rev. 61:* 183–193.

Bao, Zhiming. (1990). *On the nature of tone.* Ph.D. dissertation, MIT.

Bao, Zhiming. (1999). *The structure of Tone*. New York: Oxford University Press.

Barlow, H.B. (1959) Possible principles underlying the transformation of sensory messages. In W. Rosenblith (ed.), *Sensory Communication.* 217–234. MIT Press.

Barrie, Michael. (2003). Contrast in Cantonese vowels. *Toronto Working Papers in Linguistics, 20:* 1-19.

Barrie, Michael. (2007). Contour tones and contrast in Chinese languages. *Journal of East Asian Linguistics, 16:* 337-362.

Battistella, Edwin. (1995). Jakobson and Chomsky on Markedness. In Eva Hajičová, Miroslav Červenka, Oldřich Leška and Petr Sgall (eds.) *Prague Linguistic Circle Papers. Vol 1*. 55-72. John Benjamins.

Bauer, R. and P. Benedict. (1997). *Modern Cantonese Phonology.* Berlin: Mouton de Gruyter.

Bauer, R., Cheung, K.-H., & Cheung, P.-M. (2003). Variation and merger of the rising tones in Hong Kong Cantonese. *Language Variation and Change 15*: 211–225.

Baxter, William H. (1992). *A Handbook of Old Chinese Phonology.* Berlin: Mouton de Gruyter.

Bennett, C.H., P. Gacs, M. Li, P.M.B. Vitanyi, and W. Zurek. (1998). Information distance. *IEEE Transactions on Information Theory*, 44:4, 1407–1423.

Bermúdez-Otero, Ricardo. (2007). Diachronic phonology. In Paul de Lacy (ed.), *The Cambridge handbook of phonology*, 497-517. Cambridge: Cambridge University Press.

Bermúdez-Otero, Ricardo. 2015. Amphichronic explanation and the life cycle of phonological processes. In Patrick Honeybone & Joseph C. Salmons (eds.), *The Oxford handbook of historical phonology*, 374-399. Oxford: Oxford University Press.

Blaho, Sylvia. (2008). *The syntax of phonology: A radically substance-free approach.* Ph.D. dissertation, University of Tromsø

Bloomfield, Leonard. (1933). *Language*. New York: Holt.

Boyd, Richard N. (1985). The logician's dilemma: deductive logic, inductive inference and logical empiricism. In Wilhelm K. Essler, H. Putnam and W. Stegmüller (eds.), *Epistemology, Methodology, and Philosophy of Science: Essays in Honour of Carl G. Hempel on the Occasion of his 80th Birthday, January 8th 1985.* Springer Netherlands.

Chan, Marjorie Kit Man. (1980). *Zhong-Shan Phonology: A Synchronic and Diachronic Analysis of a Yue (Cantonese) Dialect.* Ph.D. dissertation, University of British Columbia.

Chang, Kun. (1975). Tonal developments among Chinese dialects. *Bulletin of the Institute of History and Philology 46,* 636-710.

Chao, Yuen-Ren. (1930). A system of tone letters. *Le maitre phonetique*, 45, 24–47.

Chao, Yuen-Ren. (1976). *Aspects of Chinese Sociolinguistics. Essays by Yuen-Ren Chao.* Stanford University Press.

Chater, Nick. (1999). The Search for Simplicity: A Fundamental Cognitive Principle? *The Quarterly Journal of Experimental Phonology 52A (2),* 273-302.

Chater, Nick, and Paul Vitányi. (2003). Simplicity: a unifying principle in cognitive science? *TRENDS in Cognitive Sciences 7.1:* 19-22.

Chen, Matthew. (1976). From Middle Chinese to modern Peking. *Journal of Chinese Linguistics 4, 113-277.*

Chen, Matthew. (1986). The paradox of Tianjin tone sandhi. *Proceedings of the Chicago Linguistic Society Meeting* 22: 98–114.

Chen, Matthew. (2000). *Tone sandhi*. Cambridge: Cambridge University Press

Chen, Matthew and John Newman. (1984a). From Middle Chinese to Modern Cantonese: Part I. *Journal of Chinese Linguistics 12(1), 148-194.*

Chen, Matthew and John Newman. (1984b). From Middle Chinese to Modern Cantonese: Part I. *Journal of Chinese Linguistics 12(2), 334-388.*

Chen, Matthew and John Newman. (1985). From Middle Chinese to Modern Cantonese: Part I. *Journal of Chinese Linguistics 13(1), 122-170.*

Cheng, Teresa. (1973). The phonology of Taishan. *Journal of Chinese Linguistics,* 1, 256–322.

Cheng, Chinchuan. (1973). A quantitative study of Chinese tones. *Journal of Chinese Linguistics.* 1:93–110.

Cheng, Chinchuan. (1991). Quantifying affinity among Chinese dialects. In William S-Y. Wang (ed.), *Languages and Dialects of China,* 78–112. *Journal of Chinese Linguistics* Monograph Series no.3.

Chomsky, Noam and Morris Halle. (1968). *The sound pattern of English.* New York: Harper & Row.

Clements, George N. (2003). Feature economy in sound systems. *Phonology* 20. 287–333.

Clements, G. N., and Elizabeth V. Hume (1995). The internal organization of speech sounds. In John A. Goldsmith (ed.), *Handbook of phonological theory*. 245–306. Oxford: Blackwell.

Clements, George N., Alexis Michaud and Cédric Patin. (2010). Do we need tone features? In Hume, Elizabeth, John Goldsmith and Leo Wetzels (eds.), *Tones and Features.* De Gruyter.

Conteh, Patrick, Elizabeth Cowper, Deborah James, Keren Rice and Michael Szamosi. (1983). A Reanalysis of Tone in Mende. In Jonathan Kaye, Hilda Koopman, Dominique Sportiche and André Dugas (eds.), *Current Approaches to African Linguistics* 2: 127-137. Foris Publications, Dordrecht, Holland

Desalles, J-L. (2006). A structural model of intuitive probability. *Proceedings of the seventh International Conference on Cognitive Modeling.* 86-91. Trieste, IT: Edizioni Goliardiche.

Dessalles, J-L. (2008). *La pertinence et ses origines cognitives – Nouvelles théories.* Paris: Hermes-Science Publications.

Dessalles, J-L. (2010). Emotion in good luck and bad luck: predictions from simplicity theory. *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Portland, OR: Cognitive Science Society.

Dresher, B. Elan. (1998). On contrast and redundancy. *Paper presented at the annual conference of the Canadian Linguistic Association, May, Ottawa*. Ms., Department of Linguistics, University of Toronto.

Dresher, B. Elan. (2008). The contrastive hierarchy in phonology. In B. Elan Dresher, Avery, Peter and Keren Rice (eds.), *Contrast in Phonology: Theory, Perception, Acquisition.* pp. 11-33.

Dresher, B. Elan. (2009). *The Contrastive Hierarchy in Phonology.* Cambridge: Cambridge University Press.

Dresher, B. Elan. (2011). *The Phoneme.* In Marc van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice (eds.), *The Blackwell Companion to Phonology*, pp. 2019–2048. Massachusetts: Wiley-Blackwell

Dresher, B. Elan, Glyne Piggott, and Keren Rice. (1994). Contrast in phonology: Overview. *Toronto Working Papers in Linguistics 13.1*, iii-xvii.

Duanmu, S. (1994). Against contour tone units. *Linguistic Inquiry*, 25, 555–608.

Dwyer, David (1978). What Sort of Tone Language Is Mende? *Studies in African Linguistics* 9: 167-208.

Egerod, Sören C. (1967). Dialectology. *Current Trends in Linguistics*, pp.91–129.

Falk, R. and Konold, C. (1997) Making sense of randomness: implicit encoding as a bias for judgment. *Psychol. Rev*. 104, 301–318

Gao, J. (2004). The changing sandhi rules in Tianjin dialect. In J. Lu (ed.), *Phonetic and phonological studies on Tianjin dialect: festschrift for Professor Wang Jialing's 70th birthday*. 193-247. Beijing: Beijing Institute of Technology Press.

Garner, W. R. (1970). Good patterns have few alternatives. *American Scientist, 58,* 34-42.

Garner, W. R. (1974). *The processing of information and structure.* Potomac, MD: Erlbaum.

Goldsmith, John A. (1976). *Autosegmental phonology*. Doctoral dissertation, MIT, Cambridge, Mass.

Groot, A. W. de (1931). Phonologie und Phonetik als Funktionswissenschaften. *Travaux du Cercle Linguistique de Prague 4:* 116–147.

Gruber, J. S. (1964). *The distinctive features of tone.* Unpublished manuscript.

Hale, Mark and Charles Reiss (2008). *The phonological enterprise*. Oxford: Oxford University Press.

Hall, Daniel Currie. (2007). *The role and representation of contrast in phonological theory*. Ph.D. Dissertation, University of Toronto.

Hardy, G. H. (1940). *A mathematician's apology*. Cambridge: Cambridge University Press.

Hargus, Sharon and Keren Rice (eds.) (2005). *Athabaskan Prosody.* John Benjamins.

Hochberg, J., and McAlister, E. (1953). A quantitative approach to figure ``goodness.' ' *Journal of Experimental Psychology, 46,* 361-364.

Hou, Jinyi. (1980). Pingyao fangyan de liandu biandiao [Tone Sandhi in the Pingyao dialect]. *Fangyan 1980,* 1-14.

Hou, Jing-Yi. (1989). *Studies of the Jin dialects.* Institute of the studies of languages and cultures of Asia and Africa., Tokyo University of Foreign Studies.

Hsu, Joseph H. (1987). *A study of the various stages of development and acquisition of Mandarin Chinese by children in Chinese milieu*. National Science Council Research Report, College of Foreign Languages: Fu Jen Catholic University.

Huang, Bo-Rong. (1958). Tan-tan yang-jiang-hua yu-fa de liang ge te-dian [On two characteristics of Yangjiang syntax]. *Fangyan yu Putonghua Jikan, 1*: 60-61.

Huang, Bo-Rong. (1959). Guang-dong Yang-jiang-hua wu-liang-ci de yu-fa te-dian [Syntactic characteristics of measure words in Yangjiang of Guangdong]. *Zhongguo Yuwen* 81: 128-129.

Huang, Bo-Rong. (1966). Yang-jiang-hua de ji zhong ju-shi [Several sentence constructions in Yangjiang]. *Zhongguo Yuwen* 142: 215-222.

Huang, Bo-Rong. (1986). The sound system of the Yangjiang dialect. *Computational Analysis of Asian and African Languages 26:* 53-74.

Huang, Karen. (2009). A reconstruction of Proto-Yue vowels. *Working Papers in Linguistics (Department of Linguistics, University of Hawaii)* ***40*** *(2).*

Huang, Tsan. (2004). *Language specificity in auditory perception of Chinese tones.* Ph.D. dissertation, Ohio State University.

Huang, Tsan. (2012). Cross-linguistic and inter-dialectal differences in tone perception by native speakers of three Chinese dialects and American English. *Journal of Chinese Linguistics 40.1,* 155-179.

Hung, T., "Tianjin Tone Sandhi: Towards A Unified Approach", J. of Chinese Linguistics, 15(2):274-305, 1987.

Hyman, Larry. (1976). Phonologization. In A. Juilland (ed.) *Linguistic Studies Offered to Joseph Greenberg. Vol. 2* Saratoga: Anma Libri, pp. 407-418.

Hyman, Larry. (2010). Do tones have features? In Hume, Elizabeth, John Goldsmith and Leo Wetzels (eds.), *Tones and Features.* De Gruyter.

Hyman, Larry. (2011). Tone: Is it Different? In John Goldsmith, Jason Riggle and Alan Yu (eds.), *The Handbook of Phonological Theory, 2nd Ed.* Blackwell.

Hyman, Larry, and Russel G. Schuh. (1974). Universals of Tone Rules: Evidence from West Africa. *Linguistic Inquiry, 5.1,* 81-115.

Jakobson, Roman. (1941). *Kindersprache, Aphasie, und allgemeine Lautgesetze.* Uppsala: Uppsala Universitets Arsskrift. Trans. by A. Keiler as *Child language, aphasia, and phonological universals.* The Hague: Mouton, 1968.

Jakobson, Roman, Gunnar Fant and Morris Halle. (1952). *Preliminaries to Speech Analysis*. Cambridge, Massachusetts: MIT Acoustics Laboratory.

Jones, Daniel and Kwing Tong Woo. (1912). *A Cantonese Phonetic Reader.* London: University of London Press.

Joos, Martin (ed.) (1957). *Readings in linguistics,* vol. 1. Washington, DC: American Council of Learned Societies.

Jurafsky, Daniel (1988). On the semantics of Cantonese changed tone: women, matches, and Chinese broccoli. BLS 14. 304–318.

Kam, Tak Him. (1977). Derivation by tone change in Cantonese: A preliminary survey. *Journal of Chinese Linguistics* 5:186-210.

Kei, Joseph, Veronica Smyth, Lydia K.H. So, C.C. Lau and Ken Capell. (2002). Assessing the accuracy of production of Cantonese lexical tones: a comparison between perceptual judgement and an instrumental measure. *Asia Pacific Journal of Speech, Language and Hearing*, 7:25-38.

Kingston, John (2005). The Phonetics of Athabaskan Tonogenesis. In Hargus and Rice (2005), 137-184.

Ko, Seongyeon. (2011). Vowel contrast and vowel harmony shift in the Mongolic languages. *Language Research 47*, 23-43.

Koffka, K. (1935)**.** *Principles of Gestalt Psychology*. Lund Humphries, London.

Koffka, K. (1962). *Principles of Gestalt psychology* (5th ed.). London: Routledge and Kegan Paul. Original edition: Koffka (1935).

Kollmeier, B., T. Brand, and B. T. Meyer. (2008). Perception of speech and sound**.** *Springer Handbook of Speech Processing*. 61-82. Springer.

Kolmogorov, A. N. (1963). On Tables of Random Numbers. *Sankhyā, 25.4, 369-376.*

Kolmogorov, A. N. (1968). Three Approaches to the Quantitative Definition of Information. *International Journal of Computer Mathematics, 2:* 157-168. Originally published in Russian, in *Problemy Peredachi Informatsii, 1.1:*1-7, (1965).

Kolmogorov, A. N. (1998). On Tables of Random Numbers. *Theoretical Computer Science 207*, 387-395. Original edition: Kolmogorov (1963).

Leben, W. (1973). *Suprasegmental phonology*. PhD dissertation, MIT. Distributed by Indiana University Linguistics Club.

Leben, W. (1978). The representation of tone. In Victoria A. Fromkin (ed.), *Tone: a Linguistic Survey.* Academic Press, 177-219.

Lee, Gina. (1993). *Comparative, Diachronic and Experimental Perspectives on the Interaction between Tone and the Vowel in Standard Cantonese.* Ph.D. dissertation, Ohio State University. Distributed by OSDL.

Lee, Jackson. (2012). The representation of Contour Tones in Cantonese. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society 38*.

Leyton, Michael. (1984). Perceptual organization as nested control. *Biological Cybernetics, 51*: 141-153.

Leyton, Michael. (1986a). Principles of information structure common to six levels of the human cognitive system. *Information Sciences, 38:* 1-120. Entire journal issue.

Leyton, Michael. (1986b). A theory of information structure I: General principles. *Journal of Mathematical Psychology, 30:* 103-160.

Leyton, Michael. (1986c). A theory of information structure II: A theory of perceptual organization.

*Journal of Mathematical Psychology, 30:* 257-305.

Leyton, Michael. (1992). *Symmetry, Causality, Mind.* MIT Press.

Leyton, Michael. (2001). *A Generative Theory of Shape.* Springer-Verlag.

Li, Qian and Yiya Chen. (2012). Trisyllabic Tone Sandhi in Tianjin Mandarin. *Paper presented at the Third International Symposium on Tonal Aspects of Languages (TAL 2012), May 26–29, 2012. Nanjing, China.*

Li, Charles and Thompson, Sandra A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language* 4.185-99.

Li, M. and P.M.B. Vitányi (1992). Theory of Thermodynamics of Computation, *Proc. IEEE Physics of Computation Workshop*, Dallas, Texas, USA, 42–46.

Li, M. and Vitányi, P. (1997). An Introduction to Kolmogorov Complexity and its Applications, 2nd edition, Springer-Verlag

Li, Rong (1985). Guanhua hangyuan de fenqu [Subgroupings of Mandarin dialects]. *Fangyan*, *1,* 2-5.

Li, Xingjian and Sixun Liu (1985). Tianjin fangyan de liandu biandiao [Tone Sandhi in Tianjin dialect]. *Zhonggua Yuwen, 1*, 76-80.

Liberman, Mark and Janet Pierrehumbert (1984). Intonational invariance under changes in pitch range and length. In Mark Aronoff and Richard T. Oehrle (eds.), *Language, Sound, Structure*, pp.157–233. Cambridge, Mass.: MIT Press.

Lindblom, B., Diehl, R, Park, S.-H. & Salvi, G. (2011). Sound systems are shaped by their users: the recombination of phonetic substance. In Clements, N. & Ridouane, R. (eds.), *Where Do Features Come From? Cognitive, Physical and Developmental Bases of Distinctive Speech Categories*. Amsterdam: John Benjamins. 67-97.

Liu, Y.-Z. and J. Gao. (2003). Qu-Qu liandu biandiao guize: shehui yuyanxue bianxiang [Tone 3 sandhi rule in Tianjin dialect: a sociolinguistic variable]. *Tianjin Shifan Daxue Xuebao — Shehui Kexue Ban 5:* 65-69.

Liu, Chang (2013). Just noticeable difference of tone pitch contour change for English- and Chinese-native listeners. *Journal of the Acoustical Society of America 134*.4: 3011-3020.

Mach, E. (1895). *Popular Scientific Lectures*. Trans. By Thomas J. McCormack. Chicago: The Open Court Publishing Company.

Maddieson, Ian (1978a). The frequency of tones. *UCLA Working Papers in Phonetics* 41:43–52.

Maddieson, Ian (1978b). Universals of tone. In J. Greenberg (ed.), *Universals of Human Language*. Stanford, CA: Stanford University Press, 335–366.

Maddieson, Ian. (1984). *Patterns of Sound.* CUP: Cambridge.

Martinet, Andreé. (1952). Function, structure, and sound change. *Word*, 8.1-32.

Martinet, Andre´ (1955). *Économie des changements phonétiques.* Berne: Francke.

Matthews, Stephen and Virginia Yip. (1994). *Cantonese: A Comprehensive Grammar*. London: Routledge

Matthews, Stephen and Virginia Yip. (2011). *Cantonese: A Comprehensive Grammar. 2nd Ed.* London: Routledge.

Matisoff, James A. (2003). *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction.* Berkeley: University of California Press.

McAllister, J. W. (2005). Mathematical beauty and the evolution of the standards of mathematical proof. In M. Emmer (Ed.), *The visual mind II*. Cambridge, MA: MIT Press. 15-34

McCoy, William John Jr. (1966). *Szeyap Data for a First Approximation of Proto-Cantonese [with] Appendix.* Ph.D. dissertation, Cornell University.

Mei, Tsu-lin. (1970). Tones and Prosody in Middle Chinese and The Origin of The Rising Tone. *Harvard Journal of Asiatic Studies,* Vol. 30: 86-110.

Mei, Tsu-lin. (1977). Tones and tone sandhi in sixteen century Mandarin. *Journal of Chinese Linguistics* 5:237–260.

Mielke, Jeff. (2008). *The Emergence of Distinctive Features.* Oxford: OUP.

Mok, Peggy Pik-Ki and Peggy Wai-Yi Wong (2010). Production of the merging tones in Hong Kong Cantonese: preliminary data on monosyllables. *Proceedings of the 5th international conference on speech prosody, Chicago*. 1-4. Chicago: University of Illinois at Champaign.

Mok, Peggy P.K, Donghui Zuo and Peggy W.Y. Wong. (2013). Production and perception of a sound change in progress: Tone merging in Hong Kong Cantonese. *Language Variation and Change 25:* 341-370.

Newman, John and Anand V. Raman. (1999). *Chinese historical phonology: A compendium of Beijing and Canontese pronunciations of characters and their derivations from Middle Chinese.* Lincom Europa.

Norman, Jerry (1973). Tonal development in Min. *Journal of Chinese Linguistics.* 1:222–238.

Norman, Jerry. (1988). *Chinese.* Cambridge University Press.

Odden, David. (2010). Features impinging on tone. In Hume, Elizabeth, John Goldsmith and Leo Wetzels (eds.), *Tones and Features.* De Gruyter.

Ohala, John J. (1979). Chairman's introduction to the Symposium on Phonetic Universals in Phonological Systems and their Explanation. In *Proceedings of the Ninth International Congress of Phonetic Sciences,* 184-185. Institute of Phonetics, University of Copenhagen.

Ou, Jinghua. (2012). *Tone Merger in Guangzhou Cantonese*. MPhil thesis, Hong Kong Polytechnic University.

Oxford, Will. (2015). Patterns of contrast in phonological change: Evidence from Algonquian vowel systems. *Language* 91: 308-357

Pierrehumbert, Janet. (1980). The phonology and phonetics of English intonation. Ph.D. dissertation, Massachusetts Institute of Technology.

Pothos, Emmanuel M. and Nick Chater (2002). A simplicity principle in unsupervised human categorization. *Cognitive Science* 26: 303-343.

Pulleyblank, Edwin G. (1977). *Early Middle Chinese*. ms.

Pulleyblank, Edwin G. (1978). The nature of the Middle Chinese tones and their development to Early Mandarin. *Journal of Chinese Linguistics* 6:173–203.

Pulleyblank, Douglas (1984). *Middle Chinese: A study in historical phonology*. Vancouver: University of British Columbia Press.

Pulleyblank, Douglas (1986). *Tone in Lexical Phonology*. Dordrecht: Reidel.

Raman, Anand V. (1997). *An information theoretic approach to language relatedness*. Ph.D. dissertation, Massey University.

Rice, Keren. (1995). On Vowel Place Features. *Toronto Working Papers in Linguistics, 14*: 73-115.

Rice, Keren. (2007). Markedness in phonology. In Paul de Lacy (ed.), *Markedness, reduction and preservation in phonology*. 79–97. Cambridge: Cambridge University Press.

Sagey, Elizabeth. (1986). *The representation of features and relations in nonlinear phonology*. Ph.D. Dissertation, MIT.

Saillenfest, A. & Dessalles, J-L. (2012). Role of kolmogorov complexity on interest in moral dilemma stories. In N. Miyake, D. Peebles & R. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, 947-952. Austin, TX: Cognitive Science Society.

Sapir, Edward. (1925). Sound patterns in language. *Language 1*, 37-51. Reprinted in Joos (1957), 19-25.

Saussure, Ferdinand de. (1972 [1916]). *Cours de linguistique générale. Publié par Charles Bally et Albert Sechehaye; avec la collaboration de Albert Riedlinger. Éd. critique préparée par Tullio de Mauro.* Paris: Payot.

Savoie, D. and Ladouceur R. (1995). ‘Èvaluation et modification de conceptions, erronées au sujet des loteries". *Canadian Journal of Behavioural Science 27(2) 199-213*

Shi, Feng. 1988. Shilun Tianjinhua de shengdiao jiqi bianhua — xiandai yuyinxue biji [On tones and their recent changes in Tianjin dialect — modern phonetics notes]. *Zhongguo Yuwen 1988*.5: 351-360.

Shi, Feng. (1990). *Huanyu he Dong-Tai yu de shengdiao Geju. [Tone paradigms in Chinese and Kam-Tai languages*. Ph.D. dissertation, Nankai University.

Shi, Feng and Ping Wang. 2004. Tianjinhua shengdiao de xin bianhua [New changes in Tianjin tones]. In Feng Shi and Zhongwei Shen (eds.), *The joy of research: A festschrift in honor of Professor William S.-Y. Wang on his seventieth birthday*. 176-188. Tianjin: Nankai University Press.

Shih, Chilin (1988). Tone and intonation in Mandarin. Ms., Cornell University and AT&T Laboratories.

Shih, Chilin (1991). Pitch variation across word boundary. Paper presented at the Third North American Conference on Chinese Linguistics, Cornell.

Singh, Leher and Charlene S. L. Fu (2016). A New View of Language Development: The Acquisition of Lexical Tone. *Child Development. 87.3, 834-854.*

So, Lydia K. H. (1996). Tonal Changes in Hong Kong Cantonese. *Current Issues in Language and Society*. Vol 3.2:186-189.

So, L.K.H. and Varley, R. (1991). *Cantonese Lexical Comprehension Test.* Department of Speech and Hearing Sciences: University of Hong Kong.

Solomonoff, R. J. (1978). Complexity-based induction systems: Comparisons and convergence theorems. *IEEE transactions on Information Theory,* 24(4): 422-432.

Spahr, Christopher. (2014). A contrastive hierarchical account of positional neutralization. *The Linguistic Review*, 31(3-4): 551-585.

Stahlke, Herbert (1977). Some problems with binary features for tone, in: *International Journal of American Linguistics* 43, 1–10.

Studdert-Kennedy, Michael. (1998). Introduction: the emergence of phonology. In James R. Hurford, Michael Studdert-Kennedy and Chris Knight (eds.), *Approaches to the Evolution of Language,* 169-176. Cambridge: Cambridge University Press.

Tan, F. (1987). Tone Sandhi in the Tianjin Dialect. *Journal of Chinese Linguistics*, 15(2):228-246.

Trubetzkoy, Nikolaus. (1939). *Principles of Phonology*. Originally published in German (*Grundzüge der Phonologie)* as *Travaux du Cercle Linguistique de Prague 7*. Trans. By Christiane Baltaxe. Berkeley and Los Angeles: University of California Press.

Tse, Andrew (1992). *The acquisition process of Cantonese phonology: a case study*. MPhil. thesis, University of Hong Kong.

Tseng, Chiu-yu (1981). An acoustic phonetic study on tones in Mandarin Chinese. Ph.D. dissertation, Brown University.

Tsuji, Nobuhisa (1977). *Eight Yue dialects in Guangxi Province, China, and reconstruction of Proto-Yue phonology*. PhD Dissertation, Cornell University.

Tsuji, Nobuhisa (1980). *Comparative Phonology of Guangxi Yue Dialects.* Kazama Shobo Publishing Co., Tokyo.

Trubetzkoy, Nikolai S. (1969). *Grundzüge der Phonologie* [Principles of Phonology]. Translated by Christiane A. M. Baltaxe. Berkeley: University of California Press. Original edition: *Travaux du cercle linguistique de Prague 7, 1939.*

van der Helm, Peter A (to appear). Simplicity in perceptual organization. Johan Wagemans (Ed.) *Oxford Handbook of Perceptual Organization.* Oxford University Press.

Wan, I-Ping and Jeri Jaeger (1998). Speech errors and the represeantation of tone in Mandarin Chinese. *Phonology, 15,* 417–461.

Wang, William (1967). Phonological features of tone. *International Journal of American Linguistics*, 33, 93–105.

Wee, Lian-Hee (2010). A percolative account of Tianjin tone sandhi. *Language and Linguistics 11.1,* 21-64.

Whitaker, K. P. K. (1955). A study on the modified tones in spoken Cantonese. *Asia Major*, ser 2, 5:9-36, 184-207.

Wong, Maurice Kuen-shing. (1982). *Tone Change in Cantonese.* Ph.D. dissertation, University of Illinois at Urbana-Champaign.

Wurm, Stephen A. Rong Li, Theo Baumann, and Mei W. Lee (1987). *Language Atlas of China*, Longman.

Xu, Yi (1994). Production and perception of coarticulated tones. *Journal of the Acoustical Society of America*, *95.4*: 2240-2253.

Yip, Moira (1980). *The tonal phonology of Chinese*. Ph.D. Dissertation, MIT.

Yip, Moira (1989). Contour tones. *Phonology, 6*, 149–174.

Yip, M. (2001). Tonal features, tonal inventories and phonetic targets. *University College London Working Papers in Linguistics*, 13: 161–188.

Yip, Moira (2002). *Tone.* Cambridge University Press: NY.

You, Ru-jie. (1992). *Hanyu Fangyanxue Daolun* [*An Introduction to Chinese Dialectology*]. Shanghai: Jiaoyu Chubanshe.

Yu, Alan C. L. (2007). Understanding near mergers: the case of morphological tone in Cantonese. *Phonology 24*: 187-214.

Yuan, Jiahua. (1960). *Hanyu Fangyan Gaiyao* [*An Introduction to Chinese Dialectology*]. Beijing: Wenzi Gaige Chubanshe.

Yue-Hashimoto, Oi-kan. (1972a). *Studies in Yue Dialects 1: Phonology of Cantonese*. Cambridge University Press.

Yue-Hashimoto, Oi-kan. (1972b). Two features of Proto-Yue initials. *Journal of Asian and African Studies V.*

Yue-Hashimoto, Oi-kan. (1973). *A Guide to the Tai-shan Dialect*. Monograph #061855, Educational Resources Information Center, Washington, D.C.

Yue-Hashimoto, Oi-kan. (2005). *The Dancun Dialect of Taishan 台山淡村方言研究.* Language Information Sciences Research Centre, City University of Hong Kong.

Zhang, Jie (2002). *The effects of duration and sonority on contour tone distribution—A typological survey and formal analysis*. New York City: Routledge.

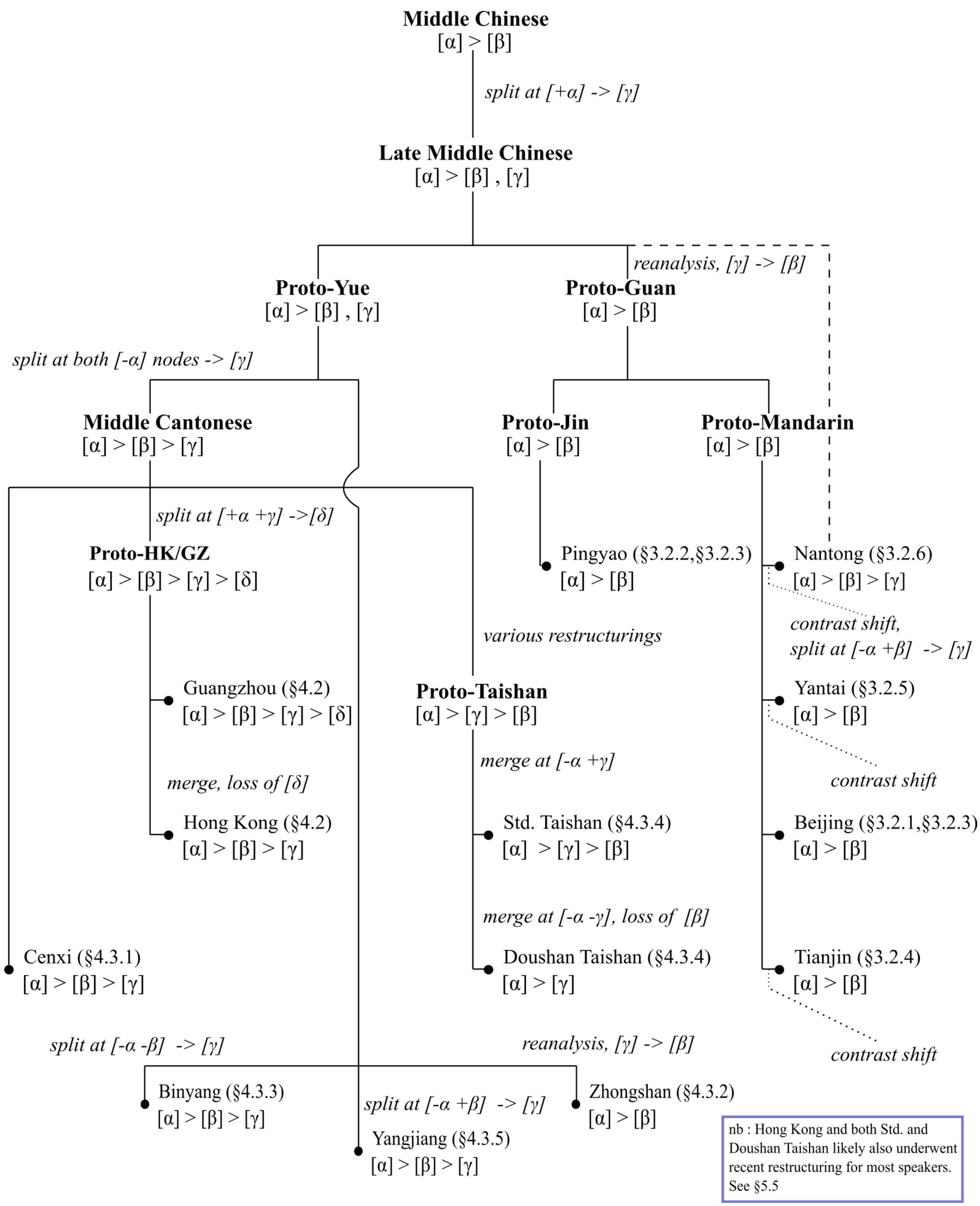
Zhang, Jie (2004). Contour tone licensing and contour tone representation. *Language and Linguistics 5.4:* 925-968.

Zhang, Jie. and Jiang Liu (2011). Tone sandhi and tonal coarticulation in Tianjin Chinese. *Phonetica, 68,* 161-191.

Zhang, Jie and Jiang Liu (2016). The productivity of variable disyllabic tone sandhi in Tianjin Chinese. *Journal of East Asian Linguistics* 25.1: 1-35

Zhang, Z-S. (1987). The Paradox of Tianjin: Another Look. *Journal of Chinese Linguistics*, 15(2):247-273.

**Appendix A:** Phylogenetic tree with changes involving loss/introduction of contrast



## **Appendix B:** List of features

[high/low] A [high] or [low] feature divides a given pitch range into two registers. Tones marked with [+high] occupy the higher register, those with [low] the lower.

[rising/falling] With pitch that contrastively rises or falls over the duration of the syllable

E.g. Cenxi (84):

| | /52/ | /21/ | /35/ | /12/ | /44/ | /22/ |
|---|---|---|---|---|---|---|
| [high] | + | - | + | - | + | - |
| [falling] | + | + | - | - | - | - |
| [rising] | - | - | + | + | - | - |

[contour] With pitch that contrastively moves over the duration of the syllable; as opposed to flat tones, which have minimal pitch motion. Directionality in [contour] tones is dependent upon an additional feature for full realization.

E.g. Zhongshan (88) (directionality determined by height)

| | /55/ | /51/ | /13/ | /22/ |
|---|---|---|---|---|
| [high] | + | + | - | - |
| [contour] | - | + | + | - |

[peripheral] A [peripheral] feature divides a given pitch range into three registers. Tones marked with [+peripheral] occupy the highest or lowest register, but not the center.

E.g. Doushan Taishan pre-restructuring (95)

| | /55/ | /11/ | /33/ | /21/ |
|---|---|---|---|---|
| [peripheral] | + | + | - | - |
| [high] | + | - | + | - |

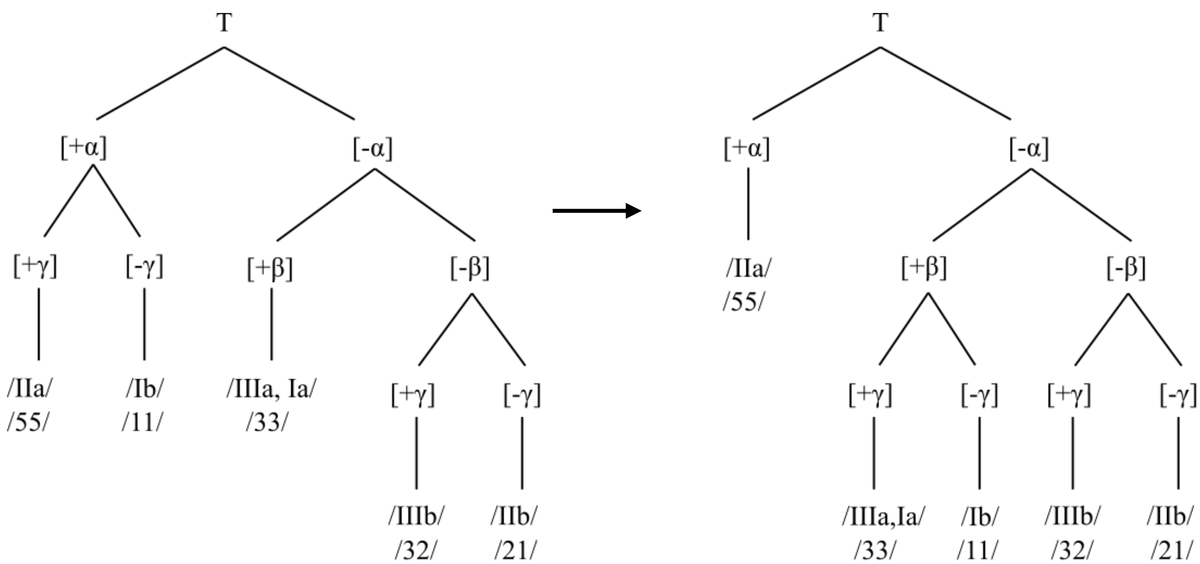
[diffusing] With pitch that contrastively moves away from the centre of a pitch range. Tones marked with [+diffusing] are high if rising or low if falling, as opposed to tones which move towards the centre of the total range, e.g. low-rising. [diffusing] is necessarily dependent upon another feature for full realization.

E.g. Proto-Tianjin (43)

| | */21/ | */45/ | */53/ | */213/ |
|---|---|---|---|---|
| [diffusing] | + | + | - | - |
| [falling] | + | - | + | - |

## Appendix C: Taishan restructuring

Standard Taishan, restructured




Doushan Taishan restructured