CHAPTER 49

TONE–MELODY MATCHING IN TONE-LANGUAGE SINGING

D. ROBERT LADD AND JAMES KIRBY

49.1 Introduction

Speakers of European languages are apt to wonder how it is possible to sing in a tone language—that is, how there can be a musical melody independent of the linguistically specified pitch on each syllable. The short answer, which looks completely uninformative at first sight, is that for a given sequence of linguistic tones, some musical melodies are a better fit than others. In this chapter we outline some of the principles that make for a good fit.

Before we start, it is important to point out that the idea of a good fit between tune and text is not unique to tone languages, though the principles are rather different in languages without lexical tone. As set out in detail in chapter 48, it is important for major stressed syllables to occur on strong musical beats in a language like English and perhaps even more important for unstressed syllables *not* to occur there. A mis-setting well known to classical musicians is Handel's treatment of the phrase *and the dead shall be raised incorruptible* in the aria 'The Trumpet Shall Sound' from *Messiah*. Handel's original version is as in (1), where the unstressed syllable *-ti-* occurs on the final strong beat of the musical phrase. (It seems plausible to suggest that Handel's native German, in which that same syllable would have borne the main stress, was responsible for the apparent musical lapse.)

(1)

Although the result is unlikely to affect the intelligibility of the text, it still feels wrong to any native speaker of English, and in most editions and most performances the setting is edited to (2).

(2)

From the point of view of musical aesthetics, the edited version arguably sounds inappropriately bouncy, yet the change is all but forced on editors and performers by the structural principles governing the match between tune and text; see further chapter 48 for metrically governed text-setting.

There is now a considerable body of work showing that structural principles—principles that are not reducible to matters of aesthetics—are involved in defining a good match between linguistic tone and musical melody as well. A particularly clear demonstration of systematic correspondence between tone and melody is found in a paper by Chan (1987b) on Cantonese pop songs. Chan observed that the sequences of syllable tones in lyrical stanzas sung to the same melody were strikingly similar from verse to verse, even though the words themselves were different. She treated this as evidence of a systematic relationship between a sequence of musical notes and the corresponding sequence of syllable tones in the accompanying text. She found that an important aspect of this systematic relationship involves *matching the direction of pitch change* from one syllable (or one musical note) to the next.

The critical role of the relation between one note/tone and the next seems to have been overlooked by many earlier investigators. For example, a short paper by Bright (1957) on song in Lushai (a Tibeto-Burman language of Assam State, India) compares the pitch contours of *individual musical notes* with those of the corresponding spoken syllables, finds no systematic relation, and concludes that tone is simply ignored in singing. The first author's work on Dinka song (discussed further in §49.3.4) originally made the same mistake, looking in vain for cues to linguistic tone in the acoustic properties of individual notes. This is not to assert that lexical tone never influences musical pitch—in fact, Chan (1987c) herself found such effects in songs from popular Mandarin films of the 1970s and 1980s, and more recently Schellenberg and Gick (2020) have shown experimentally that rising tones have 'microtonal' effects on the realization of individual notes in Cantonese. Rather, the point here is simply to emphasize that the relative pitch of adjacent syllables is apparently salient to both composers and listeners. Once we are alert to the importance of the relation between adjacent notes, we discover that many unrelated and typologically dissimilar tone languages with widely divergent musical traditions follow remarkably similar principles of matching tone and musical melody.

The primary goal of this chapter is to summarize recent work on tone-melody matching in tone languages and to show how focusing on pitch direction across notes makes it possible to formulate clear, tractable research questions. Our emphasis is on general structural principles, not details of performance practices or specific genres, though a clearer understanding of the structural principles should eventually make it easier to interpret performance practices as well.[1] We also do not speculate about the possibility of an overarching

[1] Tone–melody correspondence in traditional art forms, such as Cantonese opera, is also well studied (e.g. Yung 1983a, 1983b), but the problem of matching tune and text in these cases is to some extent a matter of performance practices rather than text-setting. Roughly speaking, in the acculturated (i.e. Western-influenced) musics of much of East and South East Asia, melodies are relatively fixed and texts must be chosen to fit, whereas in many 'traditional' forms the melodies are fairly abstract templates and may be modified in performance to achieve optimal tone–melody correspondence with a particular text.

theory of text-setting that would unify the formal principles of tone–melody matching in tone languages with those of traditional European metrics, but this also seems an interesting long-term goal (see §49.4).

49.2 Defining and investigating tone–melody matching

To provide a descriptive framework for discussing pitch direction across notes, we repurpose some standard terms from classical Western music theory. In polyphonic music, 'contrary motion' is present when one melodic line moves down in pitch while another is moving up; in 'similar motion', two lines move together in the same direction (without necessarily maintaining the same intervals between them); in 'oblique motion', one line moves (up or down) while the other stays on the same pitch. We propose to use the same terms to refer to the relationship between the movement of the musical melody and the corresponding linguistic pitch; for most of this chapter we refer only to the pitch direction across *pairs* of consecutive notes or syllables, which we refer to henceforth as 'bigrams'. Given a tonal sequence /tàpá/, where grave accent represents low tone and acute represents high, the bigrams in (3) exemplify the three possible ways of setting a bigram to a musical melody.[2]

(3)

a. b. c. d.

a.	b.	c.	d.
tà pá	tà pá	tà pá	tà pà
contrary setting	*similar setting*	*oblique setting I*	*oblique setting II*

It can be seen that our definitions identify two subtypes of oblique setting, which may eventually need to be distinguished more carefully. In Type I (3c), two consecutive syllables have different tones but are sung on the same note; in Type II (3d), two consecutive syllables have the same tone but are sung on different notes.

The basic constraint on tone–melody matching found in most of the tone-language song traditions that have been investigated over the past few decades might be summed up as 'avoid contrary settings': if the linguistic pitch goes up from one syllable to the next, the corresponding musical melody should not go down, and vice versa. Expressing the constraint in this way allows us to formulate hypotheses that can be tested against quantitative data. Specifically, if we tally all the bigram settings in a song corpus and label them as similar,

This issue also arises in the analysis of tone–melody correspondence in a number of South East Asian vocal traditions; see e.g. Williamson (1981) on Burmese, Tanese-Ito (1988) and Swangviboonpong (2004) on Thai, Chapman (2001) on Lao, Norton (2009) on Vietnamese, Lissoir (2016) on Tai Dam, and Karlsson (2018) on Kammu.

[2] Schellenberg (2009), who considered the general problem of tone–melody correspondence in more or less the way discussed here, refers to 'opposing', 'non-opposing', and 'parallel' pairs of notes for what we are calling 'contrary', 'oblique' and 'similar' settings.

oblique, or contrary, we find that contrary settings are far less common than would be expected if tone–melody matching were unconstrained. The avoidance of contrary settings has been shown to provide a good quantitative account of tone–melody correspondences in the songs of numerous tone languages, including Zulu (Rycroft 1959), Hausa (Richards 1972), Cantonese (Chan 1987b; Ho 2006, 2010; Lo 2013), Shona (Schellenberg 2009), Thai (Ketkaew and Pittayaporn 2014, 2015), Vietnamese (Kirby and Ladd 2016b), and Tommo So (McPherson and Ryan 2018).[3] We review some of these studies in more detail in §49.3.

Given the three types of setting, 'avoid contrary settings' is quite a weak constraint; a much stronger constraint would require similar settings. Of the languages and musical genres that have been studied, Cantonese pop music seems to come closest to this stronger constraint, but it still allows oblique settings in certain contexts (see §49.3.1). In other musical traditions, oblique settings appear entirely acceptable, and even contrary settings may not be consistently avoided. Nevertheless, it is clear that speakers of many tone languages are sensitive to the constraint against contrary settings, in that they will intuitively interpret, say, a musical line that rises from one note to the next as corresponding to a rise from a lower to a higher tone. There are numerous anecdotes of Christian hymns that have been inappropriately translated into tone languages by European missionaries who were unaware of the force of this constraint (e.g. Carter-Ényì 2018).

Before we can investigate these questions empirically in terms of the properties of bigrams, we must first deal with a specific methodological question: how to characterize bigrams that involve contour tones. A sequence of a low tone and a mid tone (/11/-/33/),[4] or a high tone and a low tone (/55/-/11/), can unambiguously be treated as going up or down, respectively, so that defining the tone–melody correspondence for the bigram is unproblematic. Suppose, however, that we have a sequence of two rising tones, such as /25/-/25/. We could consider this to be a level sequence (because it involves a sequence of identical tones), so that a similar setting would be expected to involve a sequence of identical notes in the musical melody. However, we might also define the tonal bigram in terms of the overall pitch direction from the beginning of the first tone to the end of the second, in which case it involves a rise in pitch, or in terms of the very local pitch change from the *end* of the first tone to the *beginning* of the second, in which case it involves a fall. There is no obvious way to decide this a priori, but it has the virtue of being an empirical research question, because once we are aware of the existence of the basic constraint, we can investigate how such bigrams are actually treated in text-setting. Chan's (1987c) study, which devoted considerable attention to this issue, suggests that the *final* (or target) pitch of each syllable is what counts for the purposes of defining the pitch direction across a tonal bigram, at least in Cantonese; in other words, a rising tone ending high (e.g. /25/) will be treated as high, a falling tone ending mid will be treated as mid (e.g. /53/), and so on. Whether or not this

[3] Our wording here implies that oblique settings represent a lesser violation of tone–melody matching than contrary settings—a view that is also implicit in Schellenberg's terms 'opposing' and 'non-opposing' and in the workings of McPherson and Ryan's constraint-based analysis. This assumption has rightly been called into question by Proto (2016) on the basis of her findings on Fe'Fe' Bamileke, and the issue certainly deserves closer investigation.

[4] To transcribe syllable tone, here and elsewhere we use 'Chao numbers' (Chao 1930), which are still widely used in discussing East Asian tone languages. The voice range is divided into five levels from 1 (low) to 5 (high), and the tone on each syllable is indicated by a sequence of two numbers indicating the levels at the beginning and the end of the syllable.

Table 49.1 Similar, contrary, and oblique settings, as defined by the relation between the pitch direction in a sequence of two tones and the two corresponding musical notes

		Melodic sequence		
		Up	Down	Level
Tone sequence	Up	**Similar**	*Contrary*	Oblique
	Down	*Contrary*	**Similar**	Oblique
	Level	Oblique	Oblique	**Similar**

Table 49.2 Expected frequencies of similar, oblique, and contrary settings

		Melodic sequence		
		Up	Down	Level
Tone sequence	Up	**Frequent**	*Rare*	Possible
	Down	*Rare*	**Frequent**	Possible
	Level	Possible	Possible	**Frequent**

'offset principle' is valid universally is an empirical matter, but data from some languages suggest that other parameterizations may also be possible (see §49.3).

Finally, before proceeding to discuss individual cases, we introduce the matrix diagrams we will use to present quantitative data on tallies of bigrams. If in any pair of sung notes the pitch may go up, go down, or stay the same, and the pitch direction from the first syllable tone to the second—assuming answers to the kinds of definitional questions just discussed—can be up, down, or level, then there are nine possible types of tone–melody bigrams, as shown in Table 49.1. If contrary settings are regularly avoided and similar settings weakly preferred, we might expect to find the distribution of bigrams shown in Table 49.2. By filling the cells in the matrix with actual counts of bigram types, we can test and fine-tune such predictions.[5]

49.3 Some examples

Space does not permit a thorough review of the literature but only a summary of a few representative cases with which we are most familiar. Some important early works not mentioned so far include Schneider (1943) and Jones (1959) on Ewe, Chao (1956) on Mandarin,

[5] Note that in these matrices we define 'level' in terms of identical *categories*, either linguistic or musical. Two notes in sequence that count musically as, say, A flat count as level regardless of how accurately they are realized. In this we differ from the approach taken by Schellenberg (2009) in his exploration of tone–melody matching in Shona, which is otherwise quite similar. In his analysis Schellenberg defined up, down, and level in purely acoustic terms, with a very small (1.5 Hz) threshold for considering two pitches to be the same, which meant that he counted rather few level note sequences.

and List (1961) on Thai. Most work has considered tone languages spoken in Africa or East Asia; notable exceptions include Herzog (1934) on Navajo, Pugh-Kitingan (1984) on Huli, Baart (2004) on Kalam Kohistani, and Morey (2010) on Tai Phake. We do not discuss 'pitch accent languages' at all, but it seems fairly certain that in at least some such languages similar principles are at work: in Japanese, for example, an accented syllable is likely to be sung on a higher note than the preceding and/or the following syllables (Cho 2017). We direct the interested reader to Jürgen Schöpf's online bibliography (http://www.musikologie.de/Literatur.html), which, while not exhaustive, contains many other useful references not mentioned here. Schellenberg (2012) and McPherson and Ryan (2018) also provide good overviews.

49.3.1 Cantonese pop music

Several experimental and/or quantitative studies in the past two decades have investigated tone–melody matching in Cantonese pop music or 'Cantopop'. Wong and Diehl (2002), without fully developing the three-way distinction between types of setting introduced in the previous section, did a small quantitative survey of the way tonal bigrams are actually treated melodically in text-setting. For the purposes of defining pitch direction from one tone to the next, they grouped the six tones of Cantonese into three sets, high, mid, and low, as shown in Table 49.3. Defining pitch direction on this basis, they observed similar settings in over 90% of cases. They also conducted a perceptual experiment in which Cantonese listeners were presented with short sung melodies containing an ambiguous test word; the perceived identity of the word was well predicted by assuming a match between the pitch direction in the musical melody and in the inferred linguistic tone sequence.

Subsequent work by Ho (2006, 2010) and Lo (2013) proposed a modification of Wong and Diehl's (2002) three-way classification: both authors suggest separating Tone 4 (21) from Tone 6 (22), so that a Tone 6–Tone 4 bigram would be treated as a fall (from low to extra-low) rather than as level (with both tones classed as low). Both showed that this change increases the number of instances of similar settings involving these two tones. Ho made numerous other refinements to the general approach, noting, for example, that when contrary settings occur, they almost always involve bigrams that straddle syntactic (and perhaps also musical) phrase boundaries (see e.g. Ho 2006: 1419). Lo (2013) undertook an

Table 49.3 The six Cantonese tones classified in terms of overall level, for the purposes of defining pitch direction in a sequence of two tones

Tone	Phonetic description (Chao numbers)	Classification for defining tonal pitch direction: Wong and Diehl (2002)	Ho (2006) and Lo (2013)
Tone 1	high level (55)	High	High
Tone 2	mid-high rising (35)	High	High
Tone 3	mid-level (33)	Mid	Mid
Tone 4	low falling (21)	Low	Extra-low
Tone 5	low-mid rising (23)	Mid	Mid
Tone 6	low level (22)	Low	Low

analysis of 11 Cantonese pop songs, classifying more than 2,500 bigrams into the matrix shown in Table 49.4. This shows clearly that both contrary and oblique settings are generally dispreferred, although Lo did find that in one specific context—sequences of High (/55/ or /35/) tones on a descending melody line—oblique settings seemed fairly common: over 84% of the 210 level tonal bigrams set to falling melodies were High-High sequences (a similar pattern was observed by Ho). As in Ho's work, virtually all of the instances of contrary setting in Lo's corpus occur across syntactic and/or musical phrase boundaries (Lo 2013: 31).

49.3.2 Vietnamese *tân nhạc*

Kirby and Ladd (2016b) considered tone–melody matching in a corpus of Vietnamese *tân nhạc* or 'new music', a broad term covering a number of Western-influenced genres of popular song produced since the 1940s. They took an empirical approach to the issue of whether to treat tonal bigrams as level, rising, or falling by enumerating all possible groupings of tones in the language and then ranking them based on the rates of similar and contrary setting that obtain under that grouping. As with Cantopop, the resulting grouping was found to refer primarily to tonal offsets, as shown in Table 49.5.

Table 49.4 Frequencies of similar (bold), oblique (underlined), and contrary (italic) settings in a 2,500-bigram corpus from Cantonese pop songs, from Lo (2013)

		Melodic sequence		
		Up	Down	Level
Tone sequence	Up	**931**	*32*	22
	Down	*56*	**828**	39
	Level	103	210*	**444**

* Mostly with H tones.
(Lo 2013)

Table 49.5 The six Vietnamese tones classified in terms of overall level, for purposes of defining pitch direction in a sequence of two tones

Tone (traditional name)	Phonetic description (Chao numbers)	Classification for defining tonal pitch direction (Kirby and Ladd 2016b)
ngang	mid-level (33)	Mid
huyền	mid-falling (32)	Low
sắc	rising (24)	High
nặng	checked (21ˀ)	Extra-low
hỏi	low falling (21)	Extra-low
ngã	broken (3ˀ5)	High

Table 49.6 Frequencies of similar (bold), oblique (underlined), and contrary (italic) settings in a corpus from Vietnamese 'new music'

		Melodic sequence		
		Up	Down	Level
Tone sequence	Up	**1136**	*81*	<u>84</u>
	Down	*72*	**1111**	<u>59</u>
	Level	<u>256</u>	<u>273</u>	**473**

(Kirby and Ladd 2016)

As shown in Table 49.6, the majority of bigrams involved similar settings, although the overall percentage (77%) was lower than that previously found for Cantonese. Oblique settings appear to be tolerated to a certain degree, particularly when sequences of the mid-level (*ngang*) tone are involved, but contrary settings were found to be comparatively rare, again suggesting that avoidance of contrary setting may generally be more important than achieving 'parallelism' between tonal and melodic pitch direction (see also Ketkaew and Pittayaporn 2014; McPherson and Ryan 2018).

49.3.3 Contemporary Thai song

Central Thai is normally regarded as having five lexical tones, conventionally labelled High, Mid, Low, Falling, and Rising. However, these labels are somewhat misleading regarding their modern phonetic realizations, and work on tone–melody mapping in contemporary Thai song has consistently found lower rates of similar setting than in, for example, Cantonese or Vietnamese, with rates reported as low as 40% (List 1961; Ho 2006; Ketkaew and Pittayaporn 2014). As discussed in §49.2, however, these rates are partly a function of how contour tones are classified for the purposes of defining the pitch direction across a bigram. Both Ho (2006) and Ketkaew and Pittayaporn (2014) note that the *onset* of the Falling tone /42/ appears to be more important than the offset, while the *offset* of the Rising tone /24/ is more relevant for tone–melody correspondence. In other words, both Falling and Rising tones appear to function more like High tones in at least some contemporary Thai genres. This suggests that the 'offset principle', while predicting high rates of tone–melody correspondence in Vietnamese and Cantonese, may not be a universal principle of tonal text-setting.

Rather than group individual tones into discrete categories, Ketkaew and Pittayaporn (2014) treated the 25 possible types of tonal bigrams separately, grouping them into three categories (rising, falling, and level) based on the type of melodic transition they most often occurred with. For example, they treated the sequence Falling-Falling as a falling transition, but Rising-Rising as a level transition. The resulting correspondence matrix for their corpus of 30 Thai pop songs is reproduced in Table 49.7.

In addition to highlighting the problem of tonal classification, studies on Thai song raise a number of other issues. Many short syllables of polysyllabic Thai words are realized as

Table 49.7 Frequencies of similar (bold), oblique (underlined), and contrary (italic) settings in a bigram corpus from 30 Thai pop songs

		Melodic sequence		
		Up	Down	Level
Tone sequence	Up	**1091**	*317*	230
	Down	*415*	**1039**	275
	Level	426	483	**594**

(Ketkaew and Pittayaporn 2014)

toneless or with reduced tone on the surface (Potisuk et al. 1994; Peyasantiwong 1986), so simply considering all musical bigrams in an automatic fashion may inflate the number of 'true' mismatches. Other factors such as word status (lexical vs. grammatical), note value, and interval range may also need to be taken into account in building up a more accurate picture of tonal text-setting constraints in languages such as Thai.

49.3.4 Traditional Dinka songs

Some years ago the first author was involved in a large funded project[6] on Dinka song. One of the research aims was to try to understand how tone and other 'suprasegmental' phonological features are manifested in singing. This is not a peripheral topic in Dinka, for two reasons. First, song is ubiquitous in Dinka culture: individuals compose and own songs that are used for a wide range of social purposes (Deng 1973; Impey 2013). Second, Dinka phonology has an unusually rich suprasegmental component, including a two-way voice–quality distinction (Andersen 1987b), a three-way quantity distinction (Remijsen and Gilley 2008), and a tonal system that is heavily used in the ablaut-based inflectional morphology (Andersen 1993, 2002; Ladd et al. 2009a). In most dialects there seem to be three distinct tones (low, high, and falling), but some may also include a fourth (rising) tone (Andersen 1987b; Remijsen and Ladd 2008).

Musically, Dinka songs are generally performed unaccompanied. They are characterized by musical phrases of variable length, with a simple rhythmic pulse and no overarching metrical structure or any analogue to stanzas. The phrases often begin with big melodic movements and end with long stretches of syllables sung on the same pitch. The musical

[6] 'Metre and Melody in Dinka Speech and Song', funded by the UK Arts and Humanities Research Council, 2009–2012. The project was the initiative of Bert Remijsen (Edinburgh); others involved were Angela Impey (SOAS, University of London) and Miriam Meyerhoff (formerly Edinburgh, now Wellington), and, among others in South Sudan, Peter Malek and Elizabeth Achol Deng. Much of the effort of the project was devoted to assembling an archive of some 200 Dinka songs, available from the British Library (http://sounds.bl.uk/World-and-traditional-music/Dinka-songs-from-South-Sudan). The question addressed here—tone–melody matching in Dinka—is unfortunately still best described as 'work in progress'.

scale is an unremarkable anhemitonic pentatonic scale (in ordinary language, any scale that can be played using only the black notes of a piano keyboard), which is found in many unrelated musical traditions around the world. The texts are of paramount importance in many performances, which means that the question of whether and how tonal identity is conveyed in music is of considerable interest.

As mentioned in the introduction, our initial approach to this question assumed that cues to the tone on a given syllable would be found on the syllable itself—for example, that low toned syllables might be produced slightly below pitch and high toned syllables slightly above. In a good deal of exploratory work on some of the songs in our corpus, we found no such pattern.[7] The absence of any such effect for high and low tones led us to conclude that there are no local phonetic cues to the linguistic tone of a sung note.

By contrast, bigram-based effects of the sort illustrated so far are certainly present. A preliminary tally based on about one minute each of three songs sung by three different singers—a total of 334 bigrams—is displayed in Table 49.8. It can be seen that just over half the bigrams exhibit similar settings and less than 10% involve contrary settings. As just mentioned, musical phrases in Dinka songs often end with long sequences of identical notes, which makes oblique setting (of Type I; see example (3c)) very likely; this can be seen in the rightmost column of the matrix. Oblique settings of Type II are encouraged by linguistic factors: over half the syllables in running text have low tone, and it seems accurate to describe low tone as 'unmarked', which means that there are many sequences of two, three, or even more low tones. This is reflected in the bottom row of the matrix.

Fuller investigation of tone–melody mapping in Dinka is difficult for a variety of practical reasons, the most important being that tone varies somewhat from dialect to dialect and is not marked in the developing standard orthography. There are also significant questions about how to deal with toneless syllables, which are fairly common in Dinka. The tallies in Table 49.8 are based on a small set of tonally transcribed songs that we intend to use in a fuller analysis.

Table 49.8 Frequencies of similar (bold), oblique (underlined), and contrary (italic) settings in a 355-bigram pilot corpus from three Dinka songs

		Melodic sequence		
		Up	Down	Level
Tone sequence	Up	**56**	*19*	21
	Down	*6*	**50**	19
	Level	45	39	**69**

[7] The negative finding is almost certainly meaningful, because our methods were sensitive enough to detect other small phonetic effects in the same song corpus. Specifically, we readily identified effects of vowel-intrinsic fundamental frequency (f0): mean f0 values for high vowels produced on a given musical note are slightly higher than for mid vowels, which in turn are slightly higher than for low vowels.

49.4 Prospect

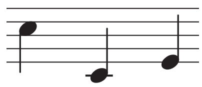
As we have suggested throughout this chapter, an important benefit of treating the problem of tone–melody correspondence in tone-language singing as a matter of text-setting constraints is that this approach allows us to formulate clear directions for further research. In this brief final section we mention a few such possibilities.

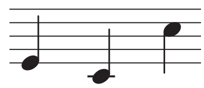
Perhaps the most obvious question that requires additional empirical evidence is one we have mentioned several times, namely how to determine whether a given tonal bigram should be expected to match a rising, level, or falling melodic sequence. We introduced this problem in connection with contour tones in §49.2 and illustrated some of the potential complexities in several of our case studies in §49.3. In Thai, for example, simply referencing the offset pitch of contour tones may prove inadequate to account for the way tonal bigrams match musical ones. Similarly, both Thai and Dinka raise the question of how to count toneless syllables in evaluating bigrams. In other languages, similar issues arise in connection with other well-known tone-language phenomena such as downstep and tone sandhi. Expressing these issues in terms of the correspondence between tonal and musical bigrams makes it possible to formulate explicit research questions with empirical answers.

A more general question concerns the variation we may expect to find from one language or musical tradition to another. As our brief case studies have sought to illustrate, languages appear to differ regarding the degree to which they actively prefer similar setting or merely seek to avoid contrary setting. There also appear to be language-specific differences with respect to oblique settings: for instance, Type II oblique settings (sequences of identical tones on a moving melody; see example (3d)) may be tolerated more with certain tones than with others. If this is the case for a particular language, for which tones might it hold, and why? To what extent can the phonetic or phonological properties of a particular tone system predict the degree to which these constraints will be violated? Much more empirical work will be necessary in order to build up a better picture of the cross-linguistic landscape.

Still another general question that needs investigation is whether it is possible and desirable to treat tone–melody mapping exclusively in terms of *local* constraints, which is the approach we have assumed throughout this chapter. This approach implies that there is no overarching plan to the melody, no sense in which long tone sequences must match whole musical phrases. Every bigram instantiates one of three types of setting, and the constraints on tone–melody correspondence are defined in terms of the type of setting on each bigram. Even the conditions under which the constraints can be relaxed (such as the fact that contrary settings may occur across syntactic and/or musical phrase boundaries) seem to be definable in strictly local terms. It is thus natural to ask whether there are instances in which sequences larger than bigrams are required, or whether highly local constraints (in conjunction with references to high-level structure, such as musical or syntactic boundaries) will suffice. To mention just one example, Carter-Ényì (2016) discusses three-note sequences like those in (4); he shows that, if there is a difference in pitch between the first and third notes, then there is a strong preference for the tones on those notes to differ: melody (4a) best matches H-L-M and melody (4b) best matches M-L-H.

(4)

a. 

b. 

Cases like (4) also raise the more general question of whether the *magnitude* of the pitch movement across a bigram—in musical terms, the interval between the two notes or syllables—is relevant to tone–melody correspondence. In this chapter we have treated bigrams simply as rising, falling, or level, without considering the difference between (for example) a small rise and a large one. However, there is plenty of reason to think that this difference may be important in some languages. For example, McPherson and Ryan (2018: 127–128) suggest that in their Tommo So song corpus, contrary settings involving two or more musical scale steps represent a more severe violation of text-setting principles than those involving only a single scale step. The degree to which such tendencies are observed in the tone languages of the world is an open question.

A final question for future research is the interaction between tonal and metrical constraints. This issue has broad formal and theoretical aspects but also requires much language-specific empirical investigation. With regard to formal issues, McPherson and Ryan (2018) note several possible parallels between tonal and metrical text-setting, such as sensitivity to phrasal phonology, and Halle (2015) has discussed the extent to which constraints on tune–text matching, in both tonal text-setting and traditional European metrics, are based on purely local rather than long-distance structural relations. As for empirical differences between languages and between musical genres, metrical and tonal constraints may interact in very specific ways. For example, the texts in the vocal traditions of many tone languages often themselves obey specific poetic constraints on which tones may appear in certain metrical positions; moreover, the vocal melodies themselves may not be completely rigid, but will be shaped to accommodate the tones of the text. The nature of this bidirectional interaction between tune and text clearly invites further study.