

# The Role of Music in Documenting Phonological Grammar: Two Case Studies from West Africa

Laura McPherson  
*Dartmouth College*

## 1 Introduction

While the connection between language and music has long been recognized by scholars and philosophers, the last two decades has seen a surge in interest in the topic. Much of this research has been in the realm of cognitive science rather than formal linguistics, with studies revealing heavy overlaps in sound structure (Lerdahl and Jackendoff 1983, Lehrdal 2001, Patel and Daniele 2002, Iversen et al. 2008, *i.a.*), syntax (Maess et al. 2001, Patel et al. 1998a, *i.a.*), and processing (Besson and Schön 2001, Zatorre et al. 2002, Schön et al. 2004, Patel et al. 1998b, Kölsch et al. 2004, *i.a.*).

The field of phonology has much to gain from looking at musical data, and yet music remains peripheral to the study of phonological grammar. This is despite the fact that other forms of artistic adaptation of language, such as poetic verse, have long informed phonological theory (Jakobson 1960, Kiparsky 1973 *et seq.*, Halle and Keyser 1969, 1971, Hayes 1988, Hayes and Kaun 1996, Hayes and Moore-Cantwell 2011, Ryan 2014, 2017, *i.a.*), especially with regards to stress systems and syllable weight.

In this paper, I consider what musical practices can tell us about phonological structure. I focus in particular on what I call **language-based music**, i.e. sung music or musical surrogate languages, though studies have also suggested an influence of spoken language phonology even on purely instrumental music (e.g. Patel 2008). I draw on two case studies from my fieldwork in West Africa: tone-tune association in Tommo So folk songs (vocal music) and surrogate speech on the Sambla balafon (ostensibly instrumental music). I demonstrate how these musical traditions provide evidence for phonological organization in spoken language and allow us to probe the interfaces between morphology, phonology, and phonetics. In addition to advancing phonological theory, I show how studies of language-based music can be a key tool in language documentation.

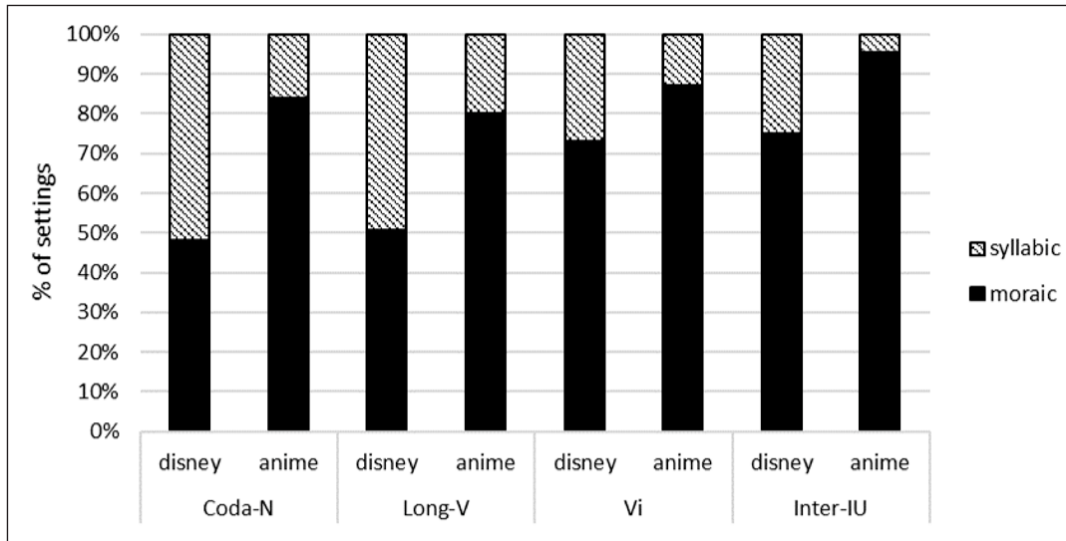
This paper is structured as follows: In §2, I briefly summarize some previous studies on music and phonological structure. I then turn to the case studies, first of Tommo So tonal textsetting in §3 and then of the Sambla balafon surrogate language in §4. §5 discusses whether differences between the two systems is a result of their modality and presents the results of a preliminary study of Seenku tonal textsetting. Finally, I conclude with suggestions for future directions in §6.

## 2 Previous studies on music and phonology

Previous work on the relationship between phonology and music has tended to focus on four main topics: textsetting, whether tonal (Herzog 1934, Leben 1983, Wong and Diehl 2002, Schellenberg 2012, Ladd and Kirby to appear, etc.) or non-tonal (Halle and Lerdahl 1993, Shih 2008, Hayes 2009, Calder 2013, Starr and Shih 2017, etc.); grouping and phrasing (Lerdahl and Jackendoff 1993, Katz and Pesetsky 2011, Katz (m.s.), etc.); rhyme (Zwicky 1976, Holtman 1996, Hanson 2003, Kawahara 2007, Katz 2015, etc.); or lastly rhythm (Patel and Daniele 2003, Seifart et al. 2018, etc.). To illustrate how data from music can advance phonological

---

\* I would like to first and foremost thank the audience at AMP for their helpful comments and lively discussion of this work. I am indebted to my musician consultants and colleagues, especially Mamadou Diabate, for teaching me about these traditions. Thanks to my research assistants, Maggie Baird, Gabriel Zuckerberg, and Lucas James for their help with the data, and to my co-author Kevin Ryan, without whom the analysis in the first case study would not be possible. This work has benefitted from the comments and insights of many audiences, including at Harvard, UCLA, Rochester, UC San Diego, the LSA, and Delaware. Finally, I gratefully acknowledge the financial support of the NSF Documenting Endangered Languages program grants BCS-0853364, BCS-1263150, and BCS-1664335.



**Figure 1:** Proportion of syllabic and moraic settings for native and translated songs.

theory, I briefly summarize one study here (Starr and Shih 2017).

Their musical study investigated the role of the mora vs. the syllable in Japanese textsetting. For decades, the mora has played an important role in the analysis of Japanese phonology, with seminal works like Vance (1987), Otake et al. (1993), Inaba (1998), among others, but the assumption remained that syllables are a universal part of human language organization and are thus still attested in Japanese. In a recent paper, however, Labrone (2012) makes the provocative claim that Japanese provides no evidence for syllables, thus calling into question their universality.

Given the central role of the syllable in studies of textsetting in other languages, Starr and Shih (2017) looked to Japanese music to determine whether the syllable played any role or if textsetting could be accounted for with the mora alone. They used a corpus of songs from animated shows and movies, including anime songs written in Japanese and Disney songs translated from English. Figure 1 summarizes the results for different mora configurations in both sets of songs.

While moraic textsetting is more common in Japanese, particularly in the native Japanese corpus, the syllable also accounts for relatively large proportions of data. As the authors argue, these settings would not be an option if the syllable were not part of Japanese phonology. This result is supported by non-musical evidence as well, for instance in the work of Kawahara (2016).

Thus, as this study shows, phonologists can turn to musical adaptation of language as a source of evidence about phonological structure. In this paper, I ask what we can learn about the phonological structure of understudied languages through the investigation of traditional musical forms.

### 3 Case study 1: Tommo So tonal textsetting

The first case study looks at vocal music in Tommo So, a Dogon language spoken by about 60,000 people in central Mali (McPherson 2013). In joint work with Kevin Ryan (McPherson and Ryan 2018), we studied a corpus of women's folk songs to determine the relationship between linguistic tone and musical melody (tonal textsetting). For background, Tommo So has two contrastive tonal primitives, H and L, which can combine to form the contour tones HL and LH on heavy syllables. Minimal pairs contrasting H and L are shown in (1):

- (1) a. /H/ *dámmá*      /LH/ *dàmmá*  
           village           hoe  
       b. /H/ *isé*         /LH/ *isé*  
           empty           dog

In addition, syllables may be surface-underspecified for tone, receiving pitch via interpolation from surround-

		musical sequence				
		<i>2+ up</i>	<i>1 up</i>	<i>same</i>	<i>1 down</i>	<i>2+ down</i>
tone	<i>up</i> (LH)	<b>40 (80%)</b>	<b>178 (62%)</b>	243 (29%)	46 (12%)	0 (0%)
sequence	<i>same</i> (LL or HH)	7 (14%)	97 (34%)	<b>404 (48%)</b>	238 (60%)	27 (61%)
	<i>down</i> (HL)	3 (6%)	10 (4%)	193 (23%)	<b>112 (28%)</b>	<b>17 (39%)</b>

**Table 1:** Relationship between tonal bigrams and musical bigrams in Tommo So vocal music

ing specified tones (McPherson 2011). Finally, like other Dogon languages, Tommo So displays a complex system of replacive grammatical tone, where lexical tone is completely overwritten with construction-specific melodies triggered by different modifiers in the noun phrase (Heath and McPherson 2013, McPherson 2014, McPherson and Heath 2016).

We set out to investigate the relationship between linguistic tone and musical melody in Tommo So women’s folk songs. There is a vast and growing literature on “tone-tune association”, or tonal textsetting; for surveys of the literature, see Schellenberg (2012), McPherson and Ryan (2018), and Ladd and Kirby (to appear). A comparison of the results of these studies shows much cross-linguistic variation in rates of tone-tune association, from languages like Cantonese where 92% of note-to-note transitions show parallel movement of tone and tune (Wong and Diehl 2002), to languages like Shona, where only 53% of transitions are parallel (Schellenberg 2009). Most languages surveyed also show tolerance for so-called “oblique” transitions, where one parameter (tone or tune) remains level and the other goes up or down. When considering both parallel and oblique transitions together (i.e. all non-opposing transitions), the rate of tone-tune association for Cantonese climbs to 98% and for Shona it climbs to 67%. Interestingly, studies have so far not revealed any areal or genetic tendencies in rates of tone-tune association; it appears to a language- and culture-specific parameter. As a case in point, while Cantonese pop music shows a tight relationship between tone and melody, Mandarin pop music shows little correspondence at all (Ho 2006).

**3.1 The data corpus** We drew our data from an hour and a half of singing that I recorded in Mali in 2012, performed by three female singers. The songs are largely call and response, where the solo verse is an elaboration of the chorus repeated by the other singers. The music is based on the following pentatonic scale, shown with its closest notes of the Western scale and numbers representing scale degrees:

(2) *Eb F Ab Bb C*  
1 2 4 5 6

In certain songs, G takes the place of Ab, and in some rare cases, the leading tone D is also found, but the five notes in (2) represent the core of the Dogon scale.

Of this hour and a half, I transcribed eleven minutes consisting of eight songs. This resulted in a corpus of 172 musical lines or 2223 musical bigrams, defined as sequences of two notes. We coded each bigram for a range of musical and linguistic factors, including: surface tone (e.g. HH, HL, etc.), whether that tone is lexical or grammatical, underlying tone (which can differ from surface tone in the case of grammatical tone), change in note (-1, 0, etc.), juncture strength (i.e. within-word, within clitic group, etc.), whether the lyrics were improvised or rote, the position of the bigram within a line, and the singer(s).

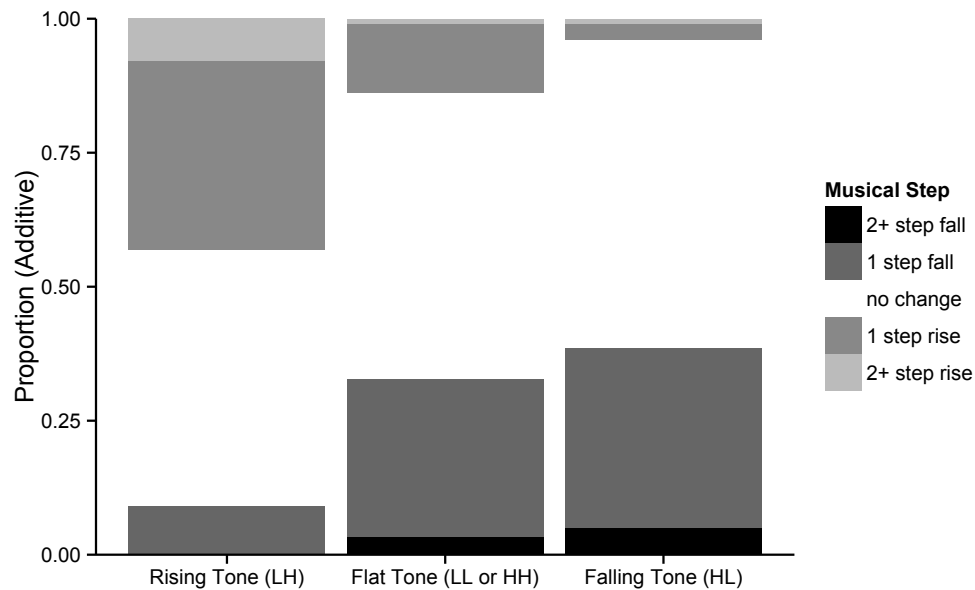
**3.2 Tone-tune association** Following the methodology in Schellenberg (2012), Ladd and Kirby (to appear), and others, we classified each bigram as one of the following:

- Parallel: e.g. rising tone with rising melody, level tone with level melody
- Contrary: e.g. rising tone with falling melody, falling tone with rising melody
- Oblique: e.g. level tone with rising melody, falling tone with level melody

The Tommo So data reveal that contrary mappings (in the gray cells) are strongly avoided, but oblique mappings (non-boldface text) are tolerated:

These results are repeated as a stacked bar graph in Figure 2.

Interestingly, cluster analysis reveals only a binary distinction in the behavior tonal transitions, between rising and non-rising tonal transitions; rising tonal transitions show avoidance of falling melodies and a fairly



**Figure 2:** Results of tonal textsetting by tone sequence and interval size

Surface tone sequence	Underlying tone sequence	% rising music	Out of N total bigrams
Nonrising	Nonrising	17%	116
Nonrising	Rising	36%	42
Rising	Nonrising	82%	11
Rising	Rising	100%	6

**Table 2:** Influence of surface and underlying tone in tonal textsetting

even split between level and rising melodies, while non-rising tonal transitions strongly avoid rising melodies and are predominantly mapped to level and falling melodies. The avoidance of contrary mappings becomes stronger with larger musical intervals.

We looked beyond the basic rate of tone-tune association in Tommo So to determine the extent to which other factors modulate this mapping. Our results show that tonal textsetting, at least in Tommo So, follows many of the same principles as non-musical verse: mapping is stricter near the ends of lines (cf. Chen 1979, Hayes 1983), stricter within words than across words (cf. Magnuson and Ryder 1970, Kiparsky 1975, Hayes et al. 2012), and stricter in rote lyrics rather than improvised lyrics. A novel finding, specific to tonal textsetting, is that lexical tone is more strictly adhered to than grammatical tone. This appears to be due to a latent effect of lexical tone, as demonstrated in the following table:

This table shows that though surface tone is the largest determinant of tonal textsetting, the underlying tone (where it differs from a surface grammatical tone melody) still exerts an influence. For instance, a word with an underlying LH tone like *giné* ‘house’ is more likely to be sung on a rising melody, even when the lyrical context places it in a construction where it would surface with a {L} grammatical overlay, such as the possessive (e.g. *Sáná<sup>L</sup> giné*, where superscript L indicates grammatical tone). I will return to this latent effect of grammatical tone below.

We modeled Tommo So tonal textsetting in Maximum Entropy Harmonic Grammar (Goldwater and Johnson 2003, Hayes and Wilson 2008, *i.a.*). The grammar contains both a general textsetting constraint, \*CONTRARY (assign a penalty to contrary mappings), as well as versions of \*CONTRARY tagged specifically for grammatical tone, sequences within the clitic group, etc., which form a stringency hierarchy with the more general constraint; see McPherson and Ryan (2018) for further discussion and exemplification.

**3.3** *What can tonal textsetting teach us about Tommo So phonology?* While the textsetting results are interesting in their own right and are suggestive of universal tendencies in artistic adaption—that is, tendencies that go beyond the bounds of metrical verse or tonal textsetting specifically—the question remains: What insights do the textsetting results provide into Tommo So phonology? I highlight two possibilities here.

First, textsetting data revealed a binary distinction between rising and non-rising tone sequences, despite the fact that the tone inventory predicts three categories: rising, level, and falling. Inclusion of a constraint targeting this binary definition of textsetting improved model fit (the AIC improved by 6). Looking back to the spoken language, it turns out that rising vs. non-rising may be an organizing principle of Tommo So tonal phonology more broadly. First, native vocabulary shows primarily two lexical tone melodies: /LH/ (rising) and /H/ (non-rising); falling melodies /HL/ make up only 6% of nouns, most of which are loanwords from Fulfulde. Further, the rising vs. non-rising distinction may play a role in grammatical tone overlays, where only non-rising melodies ({H}, {L}, {HL}) are attested.

While this may suggest a structural/phonological explanation for the binary distinction between rising and non-rising textsetting, an alternative explanation may lie in the phonetic realization of Tommo So tone. Both HH and HL sequences are phonetically falling (approximately .6 semitones for HH and 3.5 semitones for HL; McPherson 2011). The phonologically rising sequence LH, however, can be phonetically rising or in some environments even level, due to near total downdrift in the language (i.e. the last H in HLH is pronounced at the same level as the preceding L).<sup>1</sup> In other words, rising pitch is salient to Tommo So speakers, found only in a single phonological configuration, whereas level and falling are unremarkable. This may lead singers to favor rising melodies in just the single context of rising tone, whereas level and falling melodies are generally less constrained.

Thus, the binary division in Tommo So textsetting could arise from two distinct facts about the language's structure: It could either reflect singers' inherent knowledge of phonetic realization or it could reflect a deeper categorical divide between rising tonal melodies and non-rising ones, supported by both lexical and grammatical tone distribution.

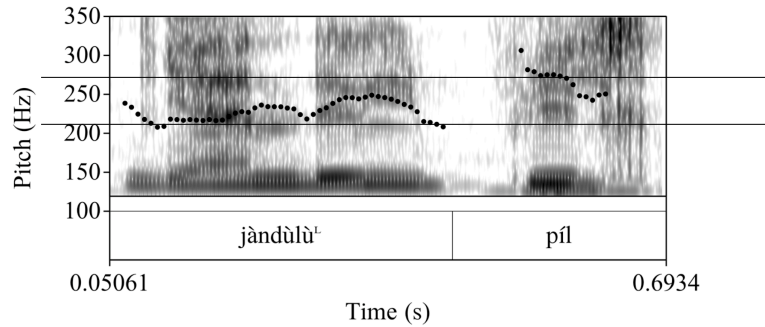
The second interesting insight from textsetting is the latent effect of underlying lexical tone in contexts where grammatical overlays have applied. In McPherson and Ryan (2018), we model this effect by having both the surface tone and the underlying tone as part of the input. The grammar contains both a general textsetting constraint \*CONTRARY as well as a version specific to lexical tone \*CONTRARY<sub>lex</sub>, which form a stringency hierarchy. In cases where surface tone and lexical tone differ, the lexical tone can still exert an influence in how the tones are set to music.

The textsetting results led me to ask whether we can find evidence of simultaneous activation of underlying lexical tone and surface grammatical tone in regular speech as well. Recall that grammatical tone in Tommo So consists primarily of replacive overlays that completely overwrite a word or stem's lexical tone. This means that lexical tone contrasts are, in principle, neutralized in grammatical tone environments, as shown in (3) with a demonstrative, which triggers a {L} grammatical overlay on the c-commanded noun (McPherson 2014, McPherson and Heath 2016):

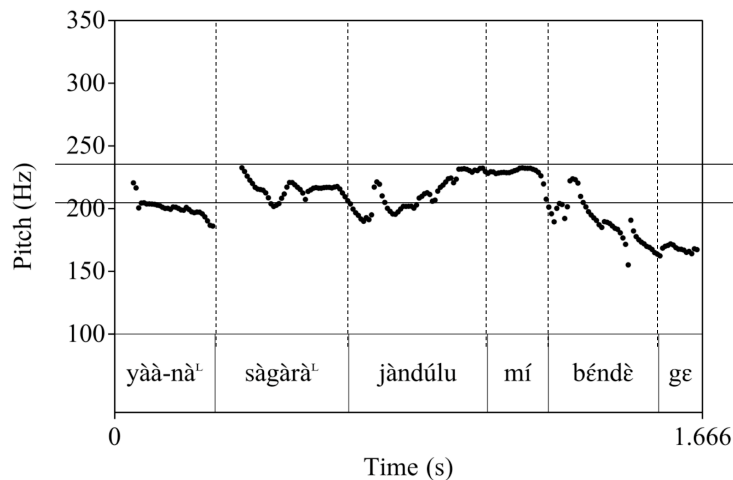
- (3) a. *dámmá*  
       'village'
- b. *dàmmá*  
       'hoe'
- c. *dàmmà' nó*  
       'this village/hoe'

Unfortunately, I do not have extensive recorded data of grammatical overlays to systematically test whether this neutralization is complete or incomplete. Nevertheless, investigation of available recordings suggests that it may be incomplete. As an illustration, consider the f<sub>0</sub> trace in Figure 3 for the phrase *jàndùlù' pìlu* 'white donkey'. The noun 'donkey' is underlyingly LH-toned (/jàndùl/, with a toneless epenthetic [u] at the end). If the {L} overlay were totally applied, then the pitch should remain flat and low throughout the noun, only

<sup>1</sup> Unfortunately, we were unable to test whether downdrift plays a role in tonal textsetting. The process appears to be phonological phrase-bounded, but how phonological phrases are determined remains an open question.



**Figure 3:** f0 trace of *jàndùlù' pílu* ‘white donkey’; lines show level of initial L tone and final H tone.



**Figure 4:** f0 trace of incomplete application of relative clause grammatical tone; lines show level of initial L and first H tone.

raising on *pílu*. Instead, Figure 3 shows that the syllable *dùl*—the underlying H-toned syllable—is pronounced at a pitch intermediate between the L tone of *jàn* and the H tone of *pílu*.

In this case, the intermediate pitch could be viewed as interpolation between the preceding L and the following H, rather than as an effect of underlying lexical tone. In other cases, though, the interpolation argument is unavailable, and yet we still see what looks like the effect of underlying tone. This is shown in Figure 4, an f0 trace of the following relative clause:

- (4) *yàà-nà'*      *sàgàrà'* *jàndùlu mí* *bénd-è=ge*  
 woman-HUM.SG young donkey 1SG hit-REL.PFV=DEF  
 ‘the young woman whose donkey that I hit’ (cf. *yàà-nà'* *sàgàrà'* ‘young woman’)

The figure once again shows a line from the level of the initial L tone and another line marking the level of the H tone of *jàndùlu*. We can see that both *yàà-nà'* and the initial L-tone of *jàndùlu* are pronounced at the same level, but interestingly, the grammatically lowered tone of *sàgàrà'* (lexically *sàgàrà'*) is pronounced at a higher level than the surrounding L tones.

I speculate that when a speaker is going to utter a sentence with a grammatical overlay, both the lexical and grammatical tone are activated; the grammatical tone is the intended target, but the underlying lexical tone may “bleed through”, producing the effects seen here. Of course, these data are preliminary, and more systematic study would be required to test whether it holds of the language more broadly.

Returning to tonal textsetting, we are once again left with the question of whether the latent effect of lexical tone in textsetting reflects subtle details of phonetic realization or whether they arise from the same

Super-high	(S)	sí	‘Terminalia (sp.)’
High	(H)	sí	‘reciprocal’
Low	(L)	sì	‘second son (proper name)’
Extra-low	(X)	sì	‘water jar’

**Table 3:** Seenku’s four tone levels

facts of mental activation.

**3.4 Interim summary** In this section, I have shown that musical melody is constrained by linguistic tone in Tommo So folk music. While a phonological analysis of the language is a prerequisite to carrying out a study of tonal textsetting, observations from textsetting also raise new questions for linguistic analysis. In the case of Tommo So, behavior of tone and musical melody may have revealed a deeper structural division between rising and non-rising tone, active in the inventory of tonal melodies, and it has also raised the possibility of incomplete neutralization of grammatical tone overlays. Both of these questions would be fertile areas of further research that deepen our understanding of the language’s phonological structure and the connections between phonetics and phonology.

In short, musical adaptation provides a window into the speaker’s implicit knowledge of their phonology that phonologists can capitalize on to direct questions of phonological theory.

## 4 Case study 2: Sambla talking balafon

The second case study broadens the definition of language-based music by focusing on a musical surrogate language—the encoding of linguistic messages in (instrumental) music. Specifically, I consider the Sambla balafon surrogate tradition from Burkina Faso. “Sambla” is an exonym for the language and ethnicity; the people themselves call the language [sɛ́ɛ-kû] (language of the Sɛ́ɛ), which I will transliterate as Seenku here, but refer to the culture and people with the term Sambla.

Seenku is a Western Mande language of the Samogo group spoken in southwestern Burkina Faso by approximately 15,000 people. It has two primary dialects, Northern and Southern; the latter has been the focus of my fieldwork since 2013 and is the subject of a forthcoming reference grammar (McPherson forthcoming). The language has a rich tone system, with four contrastive level tones, illustrated in (5):

These level tones can combine to create multiple contour tones both lexically and grammatically. These include the falling tones SX, SH, HX, and HL, the rising tones HS, LS, and XH, as well as complex contour tones like XHX and LSX.

Most Seenku vocabulary is either mono- or sesquisyllabic—a full syllable preceded by a “half” or “minor” syllable (Matisoff 1990, Pittayaporn 2015). Both monophthongs and diphthongs can be either short or long, and syllables are all open except for a coda nasal whose place is non-contrastive. Examples of typical Seenku vocabulary are illustrated in (5):

- (5) a. Monosyllabic:  
       *té* ‘who’  
       *gǔ́ɛɛ* ‘woods’  
       *dàân* ‘basket holder’
- b. Sesquisyllabic:  
       *təgê* ‘chicken’  
       *dəgǔ́ɛɛ* ‘place(s)’  
       *nəgân* ‘guinea fowl’

**4.1 Background: Surrogate languages and the Sambla balafon** Surrogate languages are defined as systems of communication in which linguistic form is mapped to non-linguistic modality. These modalities include whistling, drumming, melodic instruments, and even writing. Stern (1957) defines two types: abridgement systems, which encode phonemic aspects, and lexical ideogram systems, which symbolize concepts.

Musical surrogate languages are found worldwide on a broad range of instruments, including drums (e.g.

Scale degree:	1	b3	3	5	6	8
S	0	1	8	22	38	<b>158</b>
H	4	0	72	85	<b>154</b>	28
L	2	1	9	<b>73</b>	15	7
X	<b>101</b>	2	<b>93</b>	57	14	15

**Table 4:** Relationship between tones and notes of the scale

Yoruba, Beier 1954), flutes (e.g. Gavião, Moore and Meyer 2014), trumpets (e.g. Asante, Kaminski 2008), jaw harps (e.g. Khmu, Proschan 1994), and most importantly here, balafons (e.g. Senufo, Zemp and Soro 2010).

The balafon is a kind of resonator xylophone with many versions and variations throughout West Africa. Under each key hangs a gourd tuned to resonate at the same frequency, thus amplifying the note. Like Tommo So music, Sambla music employs a pentatonic scale, but with different intervals shown in (6):

- (6) *Sambla pentatonic scale, in the closest equivalent Western scale degrees*  
 1 b3 3 5 6

Traditional Sambla balafons have 23 keys (Strand 2009), spanning four and a half octaves, and are played by three players simultaneously. Surrogate speech is largely confined to the treble, played by the most senior musician. For further discussion of the balafon and cultural contexts of surrogacy, see McPherson (in press).

**4.2 Phonological encoding in the surrogate language** The data corpus used in this study consists of 135 phrases on the balafon transcribed for notes and inter-strike duration; the data include both elicited and naturally occurring phrases, yielding a corpus of approximately 823 words. Each of these words was coded for a number of factors, including the tone(s), the note(s) on the balafon, total duration of all strikes comprising the word, whether or not the word contains a contour tone, a long vowel, a diphthong, or a coda nasal, and finally its word shape (mono- or sesquisyllabic).

Analysis of the data reveals the following principles of encoding:

- (7) *Principles of encoding:*
1. Tone (both lexical and grammatical) is **always** encoded.
  2. Vowel length is **always** encoded.
  3. Sesquisyllabicity is **always** encoded.
  4. Segmental identity is **never** encoded.
  5. Coda nasals are **sometimes** encoded.
  6. Postlexical tone is **sometimes** encoded.

Beginning with tonal encoding, we find that the four lexical tones of the language roughly correspond to four notes of the scale. There is not a one-to-one relationship, however; exact encoding depends on the mode of the song, with S typically tied to the tonal center of the mode. In most cases, this is the octave of the tonic, shown in the following chart with 8, but other modes are also attested (e.g. centered on note 6).

With the mode centered on tones 1 and 8 of the scale (octaves) as the most general, we can see a clear pattern emerging where most S tones are played on note 8, most H on note 6, most L on note 5, and X tones split between 3 and the lower octave tonic. Interestingly, we can see in this chart that b3 is almost universally avoided. This note, known in Seenku as *jîo-bâq-dên* ‘fetish balafon key’, is said to be reserved for more spiritual uses and is seldom used in surrogate speech (Strand 2009).

To illustrate how tone encoding works, consider Figures 5-6. These illustrate how the phrases in (8-a-b) are played on the balafon, with notes of the scale on the y-axis, and words of the phrase running along the x-axis. Gray cells indicate a note strike. The central note of the mode is indicated with an arrow on the left, showing that these examples are both played in the regular “*bâq-nä*” (‘balafon mother’) mode centered on note 1 of the scale.

	S (8)				
	Sk (6)				
	T (5)				
	Bg (3)				
	J (b3)				
➤	B (1)				
	Words	á	mó	bǎ	ηέ
	Durations	181	216	207	...

Figure 5: ‘Don’t hit me!’ (NMD: July 13, 2015)

	S (8)				
	Sk (6)				
	T (5)				
	Bg (3)				
	J (b3)				
➤	B (1)				
	Words	á	mí	bǎ	ηέ
	Durations	231	235	198	...

Figure 6: ‘Don’t hit us!’ (NMD: July 13, 2015)

- (8) a. *á mó bǎ ηέ*  
 2SG 1SG.EMPH hit.IRREAL NEG  
 ‘Don’t hit me!’
- b. *á mí bǎ ηέ*  
 2SG 1PL hit.IRREAL NEG  
 ‘Don’t hit us!’

As Figures 5 and 6 show, the melody played on the balafon follows the pattern laid out by the tones of the language; H is one step below S, and X is two steps below H, just as we would expect from this four tone system.

Contour tones are encoded by playing each component tone of the contour, with the first tone played as a grace note (i.e. with a significantly shorter duration than the second tone). Figure 7 represents the balafon rendition of the following phrase:

- (9) *jô`-mərǐ nǎ mó bõ tǎgòn-tǎgòn*  
 water-drink.ANTIP.NOM PROSP 1SG.EMPH kill.IRREAL RED-completely  
 ‘I am dying of thirst.’ (Lit. ‘Thirst will kill me completely.’)

There are two contour tones in this example, first a HL contour tone on *jô`* (indicated with the dangling L tone, since the circumflex is used to represent the more common HX contour), and second a LS rising tone on the prospective auxiliary *nǎ*. Both of these contour tones are encoded on the balafon by playing a sequence of two notes, the first of which is shorter. The duration of 73 milliseconds for the L of LS-toned *nǎ* is more

	S (8)												
	Sk (6)												
	T (5)												
	Bg (3)												
	J (b3)												
➤	B (1)												
	Words	jô	mərǐ	nǎ	mó	bõ	tǎgòn	tǎgòn					
	Durations	107	152	65	217	73	182	190	193	99	188	85	...

Figure 7: ‘I’m dying of thirst’ (MD: October 14, 2016)

	S (8)					
	Sk (6)					
	T (5)					
	Bg (3)					
	J (b3)					
>	B (1)					
	Words	mó	nǎ	bǐ	sǎ	
	Duration	226	35	272	292	...

**Figure 8:** ‘I will buy goats’ MD: November 4, 2014

representative of grace note length in the corpus.

This example also illustrates the encoding of sesquisyllabic words: they are also played with two note strikes, but on the same note here, since the words *marí* and *təgòn-təgòn* have level tone. Interestingly, vowel length is encoded in precisely the same way, i.e. with two note strikes. This gives rise to ambiguity: If two note strikes are on a single note, then it could either represent a sesquisyllabic word or a long vowel; if they are on distinct notes, then it represents either a short, long, or sesquisyllabic contour tone. For further discussion and illustrations, see McPherson (in press).

**4.3 What can the balafon surrogate language teach us about Seenku phonology?** I now turn to the same question posed above for Tommo So: What can studying the balafon surrogate language tell us about the phonological structure of Seenku? First, and most simply, the balafon surrogate language provided key evidence that Seenku has four rather than three contrastive tones, as I originally believed. In my original analysis of the language, there were three contrastive levels, L, M, and H, but there were two featurally-distinct but phonetically identical M tones, one found lexically and the other derived through plural formation (McPherson 2017). At times, I would hear what sounded like upstep or downstep between two adjacent “M tones”, but I couldn’t discern a pattern. When I began eliciting phrases on the balafon, it was immediately evident that the “derived” and “lexical” M tones were consistently played differently—the derived M always lower than its supposedly lexical equivalent. Around the same time, I discovered words with this lower “derived” tone lexically (such as the numeral ‘five’ *nǎ*, always lower than ‘one’ *sóen* but higher than *sùe* ‘three’). Taken together, this led to a reanalysis of Seenku as a four-toned language, with the former M category split into L and H (and the extreme tones reclassified as X and S).

Beyond the big questions of phonemic inventory, the balafon surrogate language also provides evidence for the structure of the morphophonological component. Once again, the clearest case comes from the behavior of different kinds of tone. While in Tommo So, lexical tone and grammatical tone are treated differently for the purposes of textsetting, in Seenku, both are exceptionlessly encoded in the balafon surrogate language. Even borderline cases of grammatical tone like argument-head tone sandhi (McPherson 2019) are encoded, as can be seen in Figures 5 and 6, where ‘hit’ becomes X-toned *bǎ* after the H-toned 1sg but S-toned *bǎ* after the S-toned 1pl.

On the other hand, postlexical tone processes are only rarely encoded. The clearest example of a postlexical tone process in Seenku is contour tone simplification, whereby rising LS is simplified to either L or !S (downstepped S), depending on the context. Figure 8 shows the balafon rendition of (10), which illustrates clearly the encoding of grammatical tone and the lack of encoding of rising tone simplification:

- (10) *mó nǎ bǐ sǎ*  
 1SG.EMPH PROSP goat.PL buy.IRREAL  
 ‘I will buy goats.’

There are two cases of grammatical tone in this example: First, singular *bí* ‘goat’ is raised to S in the plural, which I argue to be a case of featural affixation (McPherson 2017). Second, this S tone then triggers *sǎ* to raise to S through argument-head tone sandhi (McPherson 2019). As Figure 8 shows, the output of both processes is represented on the balafon. However, in the spoken language, *nǎ* in this context would simplify to L-toned *nà* (via tonal absorption, Hyman and Schuh 1974). This process does not apply in the balafon surrogate language; both L and S are played as they normally would be.

These results suggest a psychologically real distinction between lexical/grammatical tone on the one hand

Scale degree:	1	b3	3	5	6	8
S	1	0	15	14	13	12
H	7	0	22	32	12	7
L	9	0	5	3	4	3
X	10	0	25	19	13	4

**Table 5:** Tone distribution by scale degree in the vocal song *Kòkò tǐ sô*

		musical sequence				
		<i>2+ up</i>	<i>1 up</i>	<i>same</i>	<i>1 down</i>	<i>2+ down</i>
tone	<i>up</i> (LH)	<b>10 (83%)</b>	<b>51 (88%)</b>	9 (14%)	1 (2%)	1 (8%)
sequence	<i>same</i> (LL or HH)	2 (17%)	7 (12%)	<b>52 (80%)</b>	15 (31%)	3 (13%)
	<i>down</i> (HL)	0	0	4 (6%)	<b>33 (67%)</b>	<b>19 (79%)</b>

**Table 6:** Relationship between tonal bigrams and musical bigrams in Seenku vocal music

and postlexical tone on the other. This distinction is available to musicians, at least subconsciously, as they encode Seenku phonological form in the balafon surrogate language. Once again, musical encoding can offer a unique glimpse into a language's phonological structure and how accessible this structure is to its speakers.

## 5 A matter of modality?

Results of Tommo So textsetting shows that singers may be sensitive to phonetic details of the language and thus are encoding a very surface level when singing. In contrast, the Sambla balafon language shows that musicians encode phonology at the lexical or morphophonological level, but crucially before postlexical tone. The two cases differ not only by language (and tonal complexity), but also by modality, with vocal music in the case of Tommo So and musical surrogate speech in the case of Seenku. In vocal music, segmental information is available in addition to tone, whereas in a musical surrogate language like the one described here, listeners can rely only on tonal and rhythmic information to decipher the message. Could this difference in modality then be related to the level at which tone is encoded?

To answer the question, I conducted a pilot study on tonal textsetting in Seenku vocal music. This preliminary study looks at just one song, *Kòkò tǐ sô* 'Koko's Song'. First, I arranged the tones in the lyrics according to notes of the scale, as I did for the balafon surrogate language. The results show no clear patterns:

However, if we instead look at tonal transitions, as we did for Tommo So tonal textsetting, the system becomes clear:

We see strict directional tone-tune association. As in Tommo So, contrary mappings (in gray cells) are strongly avoided, but unlike Tommo So, oblique mappings are also heavily dispreferred. In other words, Seenku vocal music follows very strict principles of tonal textsetting, with melodies generally following the prosodic contours laid out by the linguistic tone.

Thus far, Seenku tonal textsetting looks more like Tommo So tonal textsetting than it does the encoding of tone in the surrogate language. Closer inspection reveals that the level of encoding is likewise more akin to the Tommo So case study than the balafon language. Specifically, we see a reemergence of postlexical tone processes in sung music, including contour tone simplification as well as downdrift. Figure 9 shows the second line from the song transcribed musically; dashed lines show progressive downdrift, with the ceiling for S lowering after intervening L and X tones. Additionally, the LS-toned auxiliary *sǐ* shows what appears to be contour tone simplification—it is sung on a single note at the same level as the preceding S tone, rather than showing any rising melodic trajectory.



**Figure 9:** Downdrift (represented with dotted lines) and contour tone simplification in Seenku vocal music

These results, although preliminary, suggest that tonal textsetting in sung music takes place at a higher level, nearer to the surface, than surrogate speech, rather than the differences being related to Tommo So vs. Seenku. This could be due to the fact that segmental information is available in sung music, allowing the artist to stray from underlying lexical or grammatical forms that may help the listener retrieve the message. But if this were the case, we might expect more artistic leeway in general, i.e. that singers could deviate from even the prosodic contours of the language since the consonants and vowels are also sung. Instead, it may be the case that by using the same apparatus as speaking (i.e. the vocal tract rather than a musical instrument), singers employ the same phonological grammar including phonetic implementation that they would when speaking.

More cross-linguistic data are required to test these hypotheses. This is especially true for surrogate languages, where much of the literature is approached from an ethnomusicological angle rather than a linguistic one, meaning that subtler aspects of the phonology may not be addressed. But more work is still needed for tonal textsetting as well; though it has a relatively large literature, many studies simply report statistics for rates of association without investigating the effects of phonological and grammatical factors that can help pinpoint the level at which the singers are musically encoding phonological structure. By asking more detailed questions about how linguistic structure is musically adapted, we can in turn use the results to ask more detailed questions of the phonological structure itself.

## 6 Conclusion

In this paper, I have shown how musical adaptation of language can provide a unique source of evidence for speakers' implicit knowledge of phonological structure (e.g. the mora and the syllable in Japanese, Starr and Shih 2017, or differences between lexical and postlexical tone in the balafon surrogate language), allophonic variation and postlexical processes (Seenku tonal textsetting), and even surface phonetic details (Tommo So tonal textsetting; see also Katz 2015 on the phonetic details of imperfect rhyme in hip-hop). Musicians' sensitivities to these different levels of representation can provide evidence for boundaries between components of grammar (e.g. morphonology vs. phonology vs. phonetics).

In addition to providing a unique source of data to advance phonological theory, the study of musical adaptation is also a valuable tool in language documentation and description. By failing to encode postlexical phonology, musical surrogate languages may provide a glimpse at underlying forms that can aid in phonological analysis, while their discrete notes can amplify hard-to-hear tonal contrasts. Furthermore, language-based music is a linguistic practice and thus falls well within the bounds of language documentation, which seeks to document the linguistic and communicative practices of a speech community (Himmelmann 1998). Its inclusion in language documentation serves at least two other purposes: 1. Engaging the community, who may be more interested in music than pure language, and 2. documenting traditional music, which in many cases may be even more endangered than the language itself (Marett and Barwick 2003).

Future work will look deeper into Seenku tonal textsetting, expanding the data corpus and investigating the relationship between tonal and melodic interval size. This is an important task both for what it can teach us about Seenku phonological structure and for the fact that Sambla vocal music is a disappearing tradition. Hypotheses arising from musical study about the phonetics-phonology interface or different levels of phonological grammar will also need to be corroborated with non-musical evidence in both Tommo So and Seenku. Finally, we will need a larger cross-linguistic and cross-modal study of musical adaptation. This will allow us to answer questions like, do surrogate languages typically encode underlying structure? Is textsetting of vocal music usually sensitive to surface realization? Where does an "oral" surrogate like whistled speech (e.g. Rialland 2005) fall?

As phonologists, we are uniquely suited to study musical adaptation. In exchange, it stands to open up new avenues and insights into phonological structure.

## References

- Beier, Ulli (1954). The talking drums for the Yoruba. *African Music* 1:1, 29–31.
- Besson, Mireille & Daniele Schön (2001). Comparison between language and music. Zatorre & Peretz (eds.), *The Biological Foundations of Music*, NY Academy of Sciences, NY, 232–259.
- Calder, Jeremy (2013). Theories of syllabification in Nuxalk: Hints from text-setting. Paper presented at the annual meeting of the Linguistic Society of America.
- Chen, Matthew Y. (1979). Metrical structure: Evidence from Chinese poetry. *Linguistic Inquiry* 10, 371–420.

- Goldwater, Sharon & Mark Johnson (2003). Learning OT constraint rankings using a maximum entropy model. Spenader, Jennifer, Anders Eriksson & Osten Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 111–120.
- Halle, Morris & Samuel Jay Keyser (1969). Chaucer and the study of prosody. *College English* 28, 187–219.
- Halle, Morris & Samuel Jay Keyser (1971). A theory of meter. Halle, Morris & Samuel Jay Keyser (eds.), *English stress: its form, its growth, and its role in verse*, Harper and Row, New York, 139–180.
- Hanson, Kristen (2003). Formal variation in Robert Pinsky's *The Inferno* of Dante. *Language and Literature* 12, 309–337.
- Hayes, Bruce (1983). A grid-based theory of English meter. *Linguistic Inquiry* 14, 357–393.
- Hayes, Bruce (1988). Metrics and phonological theory. Newmeyer, Frederick J. (ed.), *Linguistics: The Cambridge survey*, Cambridge University Press, Cambridge, vol. 2, 220–249.
- Hayes, Bruce (2009). Textsetting as constraint conflict. Aroui, Jean-Louis & Andy Arleo (eds.), *Towards a typology of poetic forms: form language to metrics and beyond*, Benjamins, 43–62.
- Hayes, Bruce & Abigail Kaun (1996). The role of phonological phrasing in sung and chanted verse. *The Linguistic Review* 13, 243–304.
- Hayes, Bruce & Claire Moore-Cantwell (2011). Gerard manley hopkins's sprung rhythm: corpus study and stochastic grammar. *Phonology* 28, 235–282.
- Hayes, Bruce & Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39, 379–440.
- Hayes, Bruce, Colin Wilson & Anne Shisko (2012). Maxent grammars for the metrics of Shakespeare and Milton. *Language* 88, 691–731.
- Heath, Jeffrey & Laura McPherson (2013). Tonosyntax and reference restriction in Dogon NPs. *Language* 89:2, 265–296.
- Herzog, George (1934). Speech-melody and primitive music. *Musical Quarterly* 20, 452–466.
- Himmelman, Nikolaus (1998). Documentary and descriptive linguistics. *Linguistics* 36, 161–195.
- Ho, Wing See Vincie (2006). The tone-melody interface of popular songs written in tone languages. Paper presented at the 9th international conference on music perception and cognition, Bologna, August 22–26, 2006.
- Holtman, Astrid (1996). *A generative theory of rhyme*. Ph.D. thesis, Utrecht Institute of Linguistics, Utrecht.
- Hyman, Larry & Russell Schuh (1974). Universals of tone rules: Evidence from West Africa. *Linguistic Inquiry* 5:1, 81–115.
- Inaba, Seiichiro (1998). Moras, syllables and feet in Japanese. *Language, information and computation* 106–117.
- Iversen, John, Aniruddh Patel & Kengo Ohgushi (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America* 124:4, 2263–2271.
- Jakobson, Roman (1960). Linguistics and poetics. Sebeok, Thomas A. (ed.), *Style in Language*, MIT Press, 350–377.
- Kaminski, Joseph (2008). Surrogate speech of the Asante ivory trumpeters of Ghana. *Yearbook for Traditional Music* 40, 117–135.
- Katz, Jonah (2015). Hip-hop rhymes reiterate phonological typology. *Lingua* 160, 54–74.
- Katz, Jonah & David Pesetsky (2011). The identity thesis for language and music. Ms., MIT.
- Kawahara, Shigeto (2007). Half-rhymes in Japanese rap lyrics and knowledge of similarity. *Journal of East Asian Linguistics* 16, 113–144.
- Kawahara, Shigeto (2016). Japanese has syllables: a reply to Labrune (2012). *Phonology* 33:1, 169–194.
- Kiparsky, Paul (1973). The role of linguistics in a theory of poetry. *Daedalus* 231–244.
- Kiparsky, Paul (1975). Stress, syntax, and meter. *Language* 51, 576–616.
- Kölsch, Stefan, Elisabeth Kasper, Daniela Sammler, Katrin Schulze, Thomas Gunter & Angela Friedrici (2004). Music, language, and meaning: brain signatures of semantic processing. *Nature: Neuroscience* 7:3, 302–307.
- Labrune, Laurence (2012). Questioning the universality of the syllable: evidence from Japanese. *Phonology* 29:1, 113–152.
- Ladd, D. Robert & James Kirby (In press). Tone-melody matching in tone language singing. Gussenhoven, Carlos & Aoju Chen (eds.), *The Oxford handbook of prosody*, Oxford University Press.
- Leben, William R. (1983). The correspondence between linguistic tone and musical melody. *Proceedings of the Ninth Annual Meeting of the Berkeley Linguistics Society*, 148–157.
- Lerdahl, Fred (2001). The sounds of poetry viewed as music. *Annals of the New York Academy of Sciences* .
- Lerdahl, Fred & Ray Jackendoff (1983). *A generative theory of tonal music*. MIT Press, Cambridge.
- Maess, Burkhard, Stefan Kölsch, Thomas Gunter & Angelica Friedrici (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature: Neuroscience* 4:5, 540–545.
- Magnuson, Karl & Frank G. Ryder (1970). The study of English prosody: An alternative proposal. *College English* 31, 789–820.
- Marett, Allan & Linda Barwick (2003). Endangered songs and endangered languages. Blythe, J. & R.M. Brown (eds.), *Maintaining the links: Language identity and the land. Seventh conference of the Foundation of Endangered Languages*, Foundation for Endangered Languages, 144–151.
- Matisoff, James A. (1990). Bulging monosyllables: Areal tendencies in Southeast Asian diachrony. *Proceedings of the 16th Annual Meeting of the Berkeley Linguistics Society*, 543–559.
- McPherson, Laura (2011). *Tonal underspecification and interpolation in Tommo So*. Master's thesis, UCLA.

- McPherson, Laura (2013). *A Grammar of Tommo So*. MGL 62, De Gruyter Mouton, Berlin.
- McPherson, Laura (2014). *Replacive grammatical tone in the Dogon languages*. Ph.D. thesis, UCLA, Los Angeles, CA.
- McPherson, Laura (2017). Multiple feature affixation in Seenku plural formation. *Morphology* 27:2, 217–252.
- McPherson, Laura (2019). Seenku argument-head tone sandhi: allomorph selection in a cyclic grammar. *Glossa* 4:1, p. 22.
- McPherson, Laura (Forthcoming). *A Grammar of Seenku*. De Gruyter Mouton, Berlin.
- McPherson, Laura & Jeffrey Heath (2016). Phrasal grammatical tone in the Dogon languages. *Natural Language and Linguistic Theory* 34:2, 593–639.
- McPherson, Laura (In press). The talking balafon of the Sambla: Grammatical principles and documentary implications. *Anthropological Linguistics*.
- McPherson, Laura & Kevin Ryan (2018). Tone-tune association in Tommo So (Dogon) folk songs. *Language* 94:1, 119–156.
- Moore, Denny & Julien Meyer (2014). The study of tone and related phenomena in an Amazonian tone language: Gavião of Rondônia. *Language Documentation and Conservation* 8, 613–636.
- Otake, Takashi, Giyoo Hatano, Anne Cutler & Jacques Mehler (1993). Mora or syllable? speech segmentation in Japanese. *Journal of Memory and Language* 32:2, 258–278.
- Patel, Aniruddh (2008). *Music, language, and the brain*. Oxford University Press, Oxford.
- Patel, Aniruddh & Joseph Daniele (2002). An empirical comparison of rhythm in language and music. *Cognition* 87, B35–B45.
- Patel, Aniruddh, Edward Gibson, Jennifer Ratner, Mireille Besson & Phillip J. Holcomb (1998a). Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience* 10:6, 717–733.
- Patel, Aniruddh, Isabelle Peretz, Mark Tramo & Raymonde Labreque (1998b). Processing prosodic and musical patterns: a neuropsychological investigation. *Brain and Language* 61:1, 123–144.
- Pittayaporn, Pittayawat (2015). Typologizing sesquisyllabicity. Enfield, Nick & Bernard Comrie (eds.), *Languages of Mainland Southeast Asia: The State of the Art*, De Gruyter Mouton, Berlin/Boston, 500–528.
- Proschan, Frank (1994). Khmu play languages. *Mon-Khmer Studies* 23, 43–65.
- Rialland, Annie (2005). Phonological and phonetic aspects of whistled languages. *Phonology* 22:2, 237–271.
- Ryan, Kevin (2014). Onsets contribute to syllable weight: statistical evidence from stress and meter. *Language* 90:2, 309–341.
- Ryan, Kevin (2017). The stress-weight interface in metre. *Phonology* 34:3, 581–613.
- Schellenberg, Murray (2009). Singing in a tone language: Shona. Ojo, Akinloye & Lioba Moshi (eds.), *Selected Proceedings of the 39th Annual Conference on African Linguistics*, Cascadilla Proceedings Project, Somerville, MA, 137–144.
- Schellenberg, Murray (2012). Does language determine music in tone languages? *Ethnomusicology* 56:2, 266–278.
- Schön, Daniele & Cyrille Magnea nd Mireille Besson (2004). The music of speech: music training facilitates pitch processing in both music and language. *Psychophysiology* 41, 341–349.
- Seifart, Frank, Julien Meyer, Sven Grawunder & Laure Dentel (2018). Reducing language to rhythm: Amazonian Bora drummed language exploits speech rhythm for long-distance communication. *Royal Society of Open Science* 5, p. 170354.
- Shih, Stephanie (2008). Text-setting: A (musical) analogy to poetic meter. *The linguistics of language arts: New research programs*, University of California, Berkeley, Berkeley, CA.
- Starr, Rebecca & Stephanie Shih (2017). The syllable as a prosodic unit in Japanese lexical strata: evidence from text-setting. *Glossa* 2:1, 1–34.
- Stern, Theodore (1957). Drum and whistle “languages”: an analysis of speech surrogates. *American Anthropologist* 59, 487–506.
- Strand, Julie (2009). *The Sambla xylophone: tradition and identity in Burkina Faso*. Ph.D. thesis, Wesleyan.
- Vance, Timothy (1987). *An introduction to Japanese phonology*. SUNY Press, Albany, NY.
- Wong, Patrick M. & Randy L. Diehl (2002). How can the lyrics of a song in a tone language be understood? *Psychology of Music* 30:2, 202–209.
- Zatorre, Robert, Pascal Belin & Virginia B. Penhune (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Science* 6, 37–46.
- Zemp, Hugo & Sikaman Soro (2010). Talking balafons. *African Music* 8:4, 7–24.
- Zwicky, Arnold (1976). Well, this rock and roll has got to stop. *Chicago Linguistics Society* 12, 676–697.