

Cantopop and Speech-Melody Complex *

Edwin K. C. Li

NOTE: The examples for the (text-only) PDF version of this item are available online at:

<https://www.mtosmt.org/issues/mto.21.27.1/mto.21.27.1.li.php>

KEYWORDS: Cantopop, Cantonese, speech-melody complex, anamorphosis, anamorphic listening, perception, speech, melody, music, language, category, popular music

ABSTRACT: It is generally accepted that speech and melody are distinctive perceptual categories, and that one is able to overcome perceptual ambiguity to categorize acoustic stimuli as either of the two. This article investigates native Cantonese speakers' speech-melody experience of listening to Cantonese popular songs (henceforth Cantopop songs), a relatively uncharted territory in musicological studies. It proposes a speech-melody complex that embraces native Cantonese speakers' experience of the potentialities of speech and melody before they come into being. Speech-melody complex, I argue, does not stably contain the categories of speech or melody in their full-blown, asserted form, but concerns the ongoingness of the process of categorial molding, which depends on how much contextual information the listeners value in shaping and parsing out the complex. It follows, then, that making a categorial assertion implies a breakthrough of the complex. I then complicate speech-melody complex with the concept of "anamorphosis," borrowed from the visual arts, which calls into question the signification of the perceived object by perspectival distortion. When transferred to the sonic dimension, "anamorphic listening," I suggest, is not about at which point a sonic object becomes "distorted" but is about one's processual experience of negotiating the hermeneutic values in their different *hearing-ases*. The listener engages, then, in the process of molding and remolding, creating and negating, the two enigmatic categories. Through my analysis of two Cantopop songs and interviews with native Cantonese speakers, I suggest that Cantopop may invite an anamorphic listening, and that more broadly, it is an important, yet under-explored, genre with which to theorize about the relationships between music and language.

Received January 2020

Volume 27, Number 1, March 2021
Copyright © 2021 Society for Music Theory

所謂有始者，
繁憤未發，
萌兆牙蘖，
未有形埒垠無無蠕蠕，
將欲生興而未成物類。

What is called "There was a beginning":
Pell-mell: not yet manifest;
buds beginning, sprouts emerging;

not yet having shape or outline.

Undifferentiated, wriggling, it is on the verge of desiring to be born and flourish but not yet forming things and categories (Liu 2010, 84).

Intro

[1.1] It is generally accepted that speech and melody are distinctive perceptual categories, and that humans are able to overcome perceptual ambiguity to categorize acoustic stimuli as either of the two. The best known case of this categorial distinctiveness is perhaps the “speech-to-music” illusion, which has aroused and sustained interest in cognitive neuroscience.⁽¹⁾ The illusion, presented by Diana Deutsch and others (2011), shows that recordings of certain spoken phrases, when isolated from context and repeated, can be heard as music. In other words, through repetition, constant acoustic characteristics can generate different perceptual results. Non-musical stimuli may be perceptually transformed into musical stimuli.⁽²⁾ Studies of the speech-to-music illusion demonstrate that, while our processing of music and language recruits similar cognitive and neural resources,⁽³⁾ we are able to categorize received acoustic stimuli as speech or melody because the mental representations of each are distinct (Patel 2008, 276–82).

[1.2] The “speech-to-music” illusion reveals that much of the literature studying the cross-domain processing between music and language rests on an *a priori* categorization of auditory experience. That is, certain acoustic properties are set in advance in order that prosodic and musical elements can be compared. Those properties include, but are not limited to, fundamental frequency, duration, vocal control, and rhythmic structure. To be sure, the setting-up of such terms to measure one’s processing of music and language is conducive to understanding the similarities and differences between them; what is sacrificed, however, is the complexity of the categorization process. In this regard, Deutsch and her colleagues’ discovery represents a step forward. They find that the illusion occurs without the need to alter any property of the acoustic stimuli, thereby suggesting that whether an acoustic signal is speech or melody does not depend solely upon the inherent characteristics of the signal, but also upon a handful of mysterious cognitive operations that have yet to be fully understood.⁽⁴⁾ These experiments, however, raise more questions than they answer: for instance, with their rating scale spanning from 1 (“exactly like speech”) to 5 (“exactly like singing”), how does the numerical scale translate into real-time listening experiences and the processing of speech and melody? What do the experimenters and respondents mean by “speech” or “melody”? Why are “speech” and “melody” given as distinct categories in the first place?⁽⁵⁾

[1.3] Ethnomusicologists have attempted to answer these questions by arguing that speech and melody, though being distinctive categories, are a network of interactions between a musician and their communicative context, be it immediate or distant.⁽⁶⁾ Kofi Agawu is one of the staunch advocates of this position. In analyzing Northern Ewe music, he argues against the definitive correlation between speech tones and melodic tones, and wonders if simply analyzing the correspondence and non-correspondence between speech tones and melodic tones in musics involving tonal languages (i.e., languages in which changes in the pitch of a word often, yet not necessarily, result in changes in its meaning) does justice to the musical practice. He suggests instead that one has to hear Northern Ewe music as a dynamic construct in which a musician parses a “complex interaction of parameters,” including speech tones, contour, rhythm, and so on (1988, 143). He emphasizes that ultimately, listening to speech and melody in music is a creative act within a communicative context, and “the one and only meaning is what is communicated” (141).

[1.4] Agawu’s study, however, not unlike the scientific endeavors that distinguish speech from melody, rests on the very premise that, in his own words, “music is fundamentally different from speech” (128). That is, although one might consider speech and melody to be part of a single continuum and speak, for instance, of the musicality of speech and the speech-like quality of melody, Agawu believes that ultimately the two categories are distinct. But is there a possibility that in listening to music, the categorization of speech/melody is as much a psychological or cognitive problem as it is an aesthetic one? Can the process of categorizing speech/melody sensitize

listeners to their own craftiness and creativity and become an aesthetic stimulus that sustains their repetition of listening?

[1.5] This article takes these questions as points of departure, investigating native Cantonese speakers' speech-melody experience of listening to Cantonese popular songs [粵語流行曲] (henceforth Cantopop songs), a relatively uncharted territory in musicological studies.⁽⁷⁾ It proposes a speech-melody complex (complex, from the Latin *complectere*: to embrace) that embraces native Cantonese speakers' experience of the potentialities of speech and melody before they come into *being*. Speech-melody complex, I argue, does not stably contain the categories of speech or melody in their full-blown, asserted form, but concerns the ongoingness of the process of categorial molding, which depends on how much contextual information the listeners value in shaping and parsing out the complex.⁽⁸⁾ It follows, then, that making a categorial assertion (i.e., *this is melody and that is speech*) implies a breakthrough of the complex. I then complicate the speech-melody complex with the concept of "anamorphosis," borrowed from the visual arts, which calls into question the signification of the perceived object by perspectival distortion. When transferred to the sonic dimension, "anamorphic listening," I suggest, is not about at which point a sonic object becomes "distorted" but is about one's processual experience of negotiating the hermeneutic values in their different *hearing-ases*. The listener engages, then, in the process of molding and remolding, creating and negating, the two enigmatic categories. Through my analysis of two Cantopop songs and interviews with native Cantonese speakers, I suggest that Cantopop may invite an anamorphic listening, and that more broadly, it is an important, yet under-explored, genre with which to theorize about the relationships between music and language.

[1.6] Since the idea of speech-melody complex invites descriptions of auditory process and experience, my approach to unpacking it involves both analysis and interviews. My goal is to unravel the various culturally-conditioned aspects of listening to, interpreting, and imagining speech-melody relationships and their associated contexts in Cantopop. In December 2017 and January 2018, I conducted interviews with ten native Cantonese speakers from Hong Kong via Skype and WhatsApp. Five of the interviewees have had training in Western art music for at least five years (they are able to understand concepts like scale, tonality, and chords in Western art music and have taken instrumental classes). The other five have had minimal musical training from primary and secondary schools (they are able to understand Western staff notation but nothing more than that). The former group are music graduates from the University of Hong Kong and the Chinese University of Hong Kong (3 females and 2 males), while the latter group, who are also all university graduates, consist of three friends of mine (1 female and 2 males), one acquaintance (female), and one person introduced by my friend. Their age range is 20–23. (The issues of age and gender, however, cannot be addressed in this paper due to limited space.) I did not notify them of the topic of my study beforehand, nor did I show them the music video during the interview process since this study focuses on their auditory experience.

[1.7] This article is structured as a (Canto)pop song, which progressively complicates and enriches the ideas surrounding the concept of speech-melody complex. "Verse I" introduces an ecological hermeneutic approach to listening to Cantopop; the two "Choruses" are the analysis of Don Li's [李逸朗] "Silly Woman" ["傻女"] (2015) and FAMA's [農夫] "No Boundaries in the Sea of Knowledge" ["學海無涯"] (2007) respectively; "Verse II" connects the analyses with the concept of "anamorphic listening"; the "outro" fades away with more questions to be answered in future studies.

Verse I: An Ecological Hermeneutics of Cantopop

[2.1] Cantonese is a tonal language—that is, a language in which changes in the pitch of a word can result in changes in its meaning. Yet there is hardly any Cantonese tone system which is able to characterize Cantonese speech tones in their fullest complexity. One reason for this is that native Cantonese speakers often produce and perceive Cantonese speech tones as changed tones, not as basic tones. Basic tones occur when a word is pronounced in isolation; changed tones are the attendant tones when a word appears in a particular phonological, morphological, and semantic context.⁽⁹⁾ It follows, then, that any system of Cantonese speech tones can only remain theoretical.

That said, and despite the many tone systems in Cantonese, in this article I adopt the model of “tone letters,” which is shown in **Example 1**, owing to its ability to characterize relative pitch level and change of tonal contour.⁽¹⁰⁾

[2.2] In the tone-letter system given in Example 1, there are nine tones: two level tones, two rising tones, two falling tones, and three entering tones (entering tones are tones that occur in words ending on an unreleased consonant, -p, -t, or -k).⁽¹¹⁾ The tone letters, ranging from 1 to 5 with 1 being the lowest pitch and 5 the highest, show two important aspects of a tone and its relation to other tones in the system: first, the letters represent the change of pitch contour of a character, usually the onset and the final. For example, the character 五 /ng23/ (five) has a pitch contour from 2 to 3. Second, the letters can be used for comparisons with other tones so as to distinguish one tone from another by the relative distance of pitch, a point on which I will further elaborate in the middle sections.

[2.3] Pitch being an important factor in Cantonese to distinguish meaning, one long-standing question for linguists, psychologists, and musicologists is: how do Cantonese speech tones and sung melody interact?⁽¹²⁾ Huang Zhi-hua, in *On Writing Cantonese lyrics* (2003), writes: “That the tones of the lyrics have to conform to the melody is one of the major features of the lyrics of Cantopop. . . Nowadays Hong Kong people’s ears become more demanding that they no longer allow the prevalence of mismatches in Cantonese lyrics” (103).⁽¹³⁾ James Wong (1997), renowned Cantopop lyricist, echoes Huang: “If the tones remain unchanged when the words are sung, it is correct. If the tones are heard as different tones of other words, the words have to be changed, until the tones are sung and heard correctly” (7).⁽¹⁴⁾ Tone-melody correspondence is evidently a major concern in both the composition and perception of Cantopop.⁽¹⁵⁾ It is no wonder that Wong Shek-ling (1941) describes Cantonese speech tones as “a kind of musical pitch” (9).⁽¹⁶⁾ As one of the most prolific phonologists specializing in Cantonese, he does not intend this speech-melody mapping to be merely a metaphor, but an acoustical parallelism that serves a didactic purpose. As shown in **Example 2**, Wong uses Western staff notation to map out Cantonese speech tones and their representation in a transposable “musical scale with their proper *tone marks*” (9). For Wong, speaking in Cantonese is not unlike singing in Cantonese.

[2.4] Wong’s speech-melody mapping, however, has never been well known since its publication. And even if it had been, recent studies on Cantopop would take issue with its acoustical or structural correlation between tone and melody, and emphasize the very importance of the listeners’ perceptual experience of speech-melody relationships.⁽¹⁷⁾ Ho Wing-see Vincie, for example, reconfigures the notions of “correspondence” and “mismatch” between speech and melody into “natural” and “unnatural” perceptions from a native speaker’s listening experience. She writes, “A ‘perfect match’ refers to the mapping between a melodic transition and a tonal target transition that is satisfactorily accepted by native speakers of the language, whereas a ‘mismatch’ refers to a tone-melody pairing that sounds awkward to the native ear, whether or not the string of syllables are [sic] comprehensible, ambiguous or unintelligible when set to the song” (2010, abstract). Ho supports her claim by comparing different verses set to the same melody in Eason Chan’s [陳奕迅] “Fallen Flowers on Flowing Water” [“落花流水”]. As shown in **Example 3**, Ho observes how two pairs of different tones—high level tone 都 /dou55/ and low rising tone 有 /jau23/, and mid-level tone 些 /se55/ and mid-level tone 既 /gei33/—are cast in the same musical context (i.e., a major second). Her observation reinforces linguist Marjorie Chan’s findings that tone-melody mappings in Cantopop concern not the onset, but the rime (final) of tones, which is also called “tonal target” (1987). Tonal target transition, then, is the transition between the rime of two successive speech tones. In Ho’s case, both pairs, albeit phonologically different, sound natural to a native speaker in the musical context due to the identical directionality of tonal target transition (from 5 to 3). This naturalness, she claims, occurs even if the directionality of musical interval is different from that of the tonal target transition.⁽¹⁸⁾

[2.5] *Contra* Ho, Chow Man-ying argues that the perceptual categories of naturalness and unnaturalness are “vague, if not inaccurate” (2012, 28, 126). In her experiments, Chow takes into account the pervasiveness of non-optimal tone-melody mappings in Cantopop. She emphasizes the contextual factors that may influence the perception of speech tones in songs, including semantic,

melodic, and tonal contexts. For Chow, these muddle the correlation between tone-melody mapping and textual intelligibility. For instance, she notes that even when a tonal target transition is set to its optimal interval in Cantopop songs, there may still be perceptual mismatch (78). In short, she argues for a more fluid description of tone-melody relationships.

[2.6] For all their call for fluidity and attention to a native language user's perception, Ho's and Chow's studies can be construed as working along the correspondence/non-correspondence axis of speech and melody.⁽¹⁹⁾ Their shared goal is to problematize the positive correlation between speech and melody that was given in Example 2 and to do justice to the context-dependence of a native language user's listening experience. In so doing, they come back to the importance of the correspondence/non-correspondence axis. Yet, as Agawu suggests in his study on Northern Ewe music, analyzing speech and melody solely along the correspondence/non-correspondence axis may reflect an aesthetic bias that "the music is better if tone and tune agree" (1988, 143). Although Ho and Chow would probably cast doubt on the applicability of this statement to Cantopop, I believe that one could benefit from Agawu's call to hear speech-melody as a dynamic construct in order to break loose from the analytical straightjacket that persists in Cantopop studies.

[2.7] Scholars and lay native Cantonese speakers seem to agree on the distinctiveness of speech and melody. In the context of Cantonese opera, Bell Yung claims that speaking differs from singing in "the presence or absence of musical tones" (1989, 13). In an empirical study, Zhang Ling argues persuasively that normal speaking and singing in Cantonese are differentiated mainly on the level of pitch slope and pitch register. Speaking in Cantonese, she argues, usually involves a declining intonation and higher pitch register, whereas singing in Cantonese involves an ascending intonation and lower pitch register (2016). My interviewees' answers to the differences between speech and melody are diverse. **Example 4** shows how each respondent defined speech and melody. (Notice how the musicians tended to reply with music-theoretical terms such as "intervals" and "itches," while the non-musicians mostly responded according to the function or effect of speech and/or melody.) Although all respondents believed that a difference exists between speech and melody, some of them, especially the non-musicians, emphasized the interdependence of the two categories: to understand what melody is, their answers suggest, one has to understand what speech is, and vice versa, such that one can only be defined in terms of the other.

[2.8] Such signs of semantic interdependence, or *ecology*, between speech and melody, are harbingers of a potential collapse of the watertight distinction between speech and melody in Cantopop. As we have seen, studies on tone-melody mappings in Cantopop, in spite of their account of its prodigious complexity, presuppose that speech and melody are *perceived* as discrete entities in Cantopop. There exists, in other words, a categorial distinction between speech and melody such that one can perform an analysis on the mapping process between speech tones and melodic tones. Speech and melody, in this view, have a fixed representational structure. As I will demonstrate, however, this is not always the case. The boundary between melody and speech in Cantopop is sometimes so fuzzy that one has to rethink the fundamental questions of "What is melody?" and "What is speech?" But I am not interested in the ontology of the two categories in Cantopop, but their *ecology*—how the two perceptual categories form a complex within which the categories of speech and melody are co-defined in various contexts, and thus how they are always both under construction and on the verge of collapse.⁽²⁰⁾ Interpreting speech and melody is always interpreting *with* or *in relation to* one another, and ultimately it becomes a self-reflexive critique of such an ecology from within, a critique that makes semantics a process, not a product.

[2.9] The emphasis on process rather than on product in one's experience of listening to the formation of speech and melody within a song invites a reconsideration of the relationship between speech and melody and about their associated contexts. Rather than ask *what* speech and melody respectively are, one would profit from a hermeneutic approach that asks *when* they are.⁽²¹⁾ This involves not only the process of categorizing but also contextualizing, interpreting, breaking down, thinking about and against, and putting oneself into the categories and their relations. In answering the questions "When is speech?" and "When is melody?" one does not consider what properties speech or melody stably exemplify, but what they had been doing as sound before they became speech/melody, and what they together do and will do, with and to oneself, and what one

brings to them. A speech-melody complex is thus larger than speech and melody and entails one's ecological and hermeneutic engagements with sound.

[2.10] In the following sections, I develop these arguments through two case studies: Don Li's "Silly Woman" (2015) and FAMA's "No Boundaries in the Sea of Knowledge" (2007). These songs represent different musical and linguistic possibilities in Cantopop. While Li reinvents the Cantopop classic of the 90s by using a peculiar singing (or is it speaking?) technique called the "wailing tone" [哭腔], FAMA, a two-member group made up of performers 6-wing and C Kwan, perform Cantonese hip-hop. My choice of repertoire reflects the rich inventory of musical and linguistic phenomena that are subsumed under Cantopop.

Chorus I: Don Li's "Silly Woman" (2015)

[3.1] Don Li's "Silly Woman" is a rearrangement of Cuban-Venezuelan singer-songwriter Maria Conchita Alonso's "La Loca" (1982). It was first sung in Cantonese by Priscilla Chan in 1987, and since then has become one of the classics of Cantopop. The music video of Li's version of the song went viral on YouTube upon its release in 2015. "The Encyclopedia of Virtual Communities" [香港網絡大典] (henceforth "[The Encyclopedia](#)"), a Wikipedia-like webpage encapsulating a wide range of heated topics amongst virtual communities in Hong Kong, has gathered opinions about Li's song from internet users. Some point out that the opening of the song gets slurred together so much that it is, as the Cantonese idiom goes, as though "ghosts were eating mud" [鬼食泥]. It is, the commenters on the Encyclopedia say, apt for a funeral song [大酒店之歌].⁽²²⁾ Some sarcastically thank Li for reminding them of the Hungry Ghost Festival [盂蘭節]—a Chinese counterpart of the Western Halloween—and for making them understand what "July 14th soul-calling singing technique" [七月十四喊驚唱法] is.⁽²³⁾ Priscilla Chan, on the other hand, seems to appreciate Li's rendition. She says, "Wow! I totally lost [to him]! He imbues the music with much richer emotions than I do. But I think this is a theatrical version of 'Silly Woman' . . . To be frank, I think he is entertaining and has a gimmick. I like this version."⁽²⁴⁾ Li himself responds in a self-mocking fashion: "Many have given me the number of doctors and The Samaritan Befrienders Hong Kong . . . But really, I think music should not be confined to fixed rules."⁽²⁵⁾

[3.2] The sarcastic comments revolving around "Silly Woman" are in fact the very description of the netizens' own speech-melody experiences, which partly motivated my decision to pick Li's rendition of the song as the object of my analysis. This attests to the song's "interesting-ness" (in Sianne Ngai's sense), which is to say, its aesthetic quality that invites explanations and engenders further discourse (2012, 110–73). What does it mean, for example, when one writes that the song is like "ghosts eating mud"? Why relate Li's singing to the "soul-calling singing technique"? What does the rhetorical language of a commenter tell us about their auditory experience, not least in regard to the categorial formation of speech and melody? My answer to these questions unfolds in three parts: (1) "Ghosts Eating Mud": The Wailing Beginning; (2) The Anamorphic Chorus; (3) Desperate Screaming: The Ultimate Collapse of Speech and Melody.

"Ghosts Eating Mud:" The Wailing Beginning

[3.3] "Silly Woman" tells a story about a woman expressing nostalgic feelings for her former lover. The song does not begin with an intro, but goes straight into the verse. This is not unusual in Cantopop; what is unusual is the way Li performs. His wailing tone obfuscates not only the tonality, but also the melody itself. As we listen to the beginning of **Video Example 1** (0:00–0:08), the first line reads, "這夜我又再獨對夜半無人的空氣", which roughly means "I have to be alone tonight again."⁽²⁶⁾ The first character 這, which in this instance functions as a determiner ("this"), has three possible pronunciations: /ze25/, /ze33/, and /ze23/. The subsequent character 夜 ("night") can also be pronounced in two ways: /je25/ and /je22/. We can conclude four possible tonal target transitions from these phonological variations: 5–5, 5–2, 3–5, and 3–2 (let us recall that tonal target transition is the transition between the rime of two successive speech tones).⁽²⁷⁾ Although some combinations are less commonly used, my purpose here is not to show the frequency of usage (this will become a concern in my interviews below), but the one Li performs, that is, /ze23 je22/,

involving a tonal target transition of 3–2. As **Example 5** shows, the descending motion of speech tones coincides with that of a minor second. Yet Li adds a timbral layer to the optimal correspondence between speech tones and melody by performing it as though he were soul-calling, mumbling religious scriptures to himself with his desolate body. The soliloquy creates in the listener a depressing sense of intimacy at the outset of the song, sonically gesturing a kind of murmuring to oneself rather than singing.

[3.4] Paradoxically, such speech-like sonority is unstable in view of the *non*-correspondence between the tonal target transition and the expected musical interval in the same line. Although, as shown in **Example 6**, the tonal target transition from 人 /jan21/ to 的 /dik5/ is 1–5, which can be mapped onto a perfect fifth interval, Li’s 的 sounds otherwise. Not only does he not hit a perfect fifth, he also lowers the pitch level on reaching the tonal target in the high level tone /5/, making it almost sound like /3/. This drop in pitch level within a word affords two interpretations. First, since it usually occurs when one sighs, or speaks in the state of helplessness or despair, Li seems to be speaking in a dramatic manner.⁽²⁸⁾ The internet users’ characterization of Li’s performance as “ghosts eating mud” is in this sense valid. The idiom describes an unclear utterance—one that is too soft, slurred, or tonally “incorrect” compared to everyday utterances. Second, in performing 的, Li negates the linearity of a musical interval by adding an ascending-descending pitch contour onto a single pitch level, when /dik5/ (or /dik3/), as an entering tone (a tone that ends on an unreleased consonant), cannot, by standard pronunciation, involve any change in pitch contour. Accordingly, he transforms a speech tone into dramatized melodic tones. But rather than thinking that Li maps speech tones onto a melody or vice versa, one might claim that such a binary does not exist. One might, in listening to the tug-of-war between the correspondence and non-correspondence between speech tones and musical intervals in the first line, experience the performative tension Li creates: he is at once trying to sing and not to sing, to speak and not to speak.

[3.5] Dialogues with my interviewees further challenge the correspondence/non-correspondence observations. After letting them listen to the song, I asked them to perform two tasks: first, to reproduce what the first line sounded like to them; second, to repeat the lyrics. The five subjects with musical backgrounds sang the melody perfectly. The unexpected perfection problematizes my analysis above, because, in my view, Li does not clearly articulate the melodic intervals but approximates them. Thus, I asked the musically trained if they thought the sound they had produced was identical to Li’s. Some requested for a second hearing, and responded negative, yet saying that they could relate Li’s wailing to an approximate melody. It was, they said, exceedingly difficult to replicate what Li does. On the other hand, while one non-musician nearly perfected the semi-tones in the first few tunes, almost all failed to perform the perfect fifth from 人 to 的. Some sang a major second, others a minor third. But judging what they produced may not be appropriate in this situation. Were they even trying to sing? With that point of curiosity, I asked them the same follow-up question. Surprisingly enough, three of them responded in the affirmative, though they expressed that it was not clear to them how much a leap it was from 人 to 的. One said she tried her best to imitate his wailing tone. What does this tell us? For the musically competent, although Li does not sing a melody—he wails as though he is singing—they could decide to map it onto an imaginary one. In other words, for the musicians, there was no melody whatsoever at the outset of the song, for otherwise, they would be able to replicate what Li does. That the musicians normalized Li’s performance and remolded it into something non-existent can be due, as they claimed, partly to their training in classical music, and partly to an anxiety of categorial uncertainty. They were unsure of what they heard and thus freely attributed it to one category with which they were familiar and comfortable, and perhaps one that makes sense in a Cantopop song. For the non-musicians, the soliloquy effect overshadowed the melody itself. The near monotony of Cantonese murmuring eclipsed the supposedly more dramatic musical leap. As such, to ask whether the first line is speech or melody is to under-appreciate the process of attribution, which is contingent on the listener’s background, as well as on the will of the listeners to accommodate their attention to specific facets of the sonic object. This is why the concept of speech-melody complex comes in handy in describing the potentialities of speech and melody, the process of embracing the epistemic struggle of knowing about how we know about them.

[3.6] The second task was to repeat the lyrics. Some respondents started their line by pronouncing the character 這 as /ze25/ rather than /ze23/ or /ze33/. Intrigued by this phenomenon, I asked them the rationale behind their decision. One respondent said, “/ze23/ (the pronunciation in the song) sounds archaic and pedantic, but I guess some people speak like this.”⁽²⁹⁾ Other respondents began their line by pronouncing the “unusual” /ze23/. They said, upon reflection, that although somewhat conditioned by the song, they thought that it was a fine pronunciation with regard to the literary context. The pronunciation would not affect the overall meaning. The difference in opinion among respondents adds another dimension to the relationship between speech and melody. That is, one can regard the first line as a melody if one eliminates the option of its being a speech based on its obsolete pronunciation.⁽³⁰⁾ The standard pronunciation is altered by, and thus subservient to, the musical melody. On the other hand, one can perceive that the archaic pronunciation enhances the literary effect of the speech, rendering it even more speech-like. The fact that Li restrains the leap of pitch contour (/23/ instead of /25/) also contributes to the soliloquy effect at the song’s outset. Parsing a speech-melody complex, accordingly, depends on what the listener values, and how they turn sonic signals into meaningful information.

The Anamorphic Chorus

[3.7] After the second verse, the chorus repeats itself four times. The chorus music is the same, but it alternates between two different texts, which I call Chorus 1 and Chorus 2 (2:05–3:03). The lyrics of both choruses are given in **Example 7** and **Example 8** below for readers to follow along as they listen.

[3.8] “As the music goes on, it sounds nothing like melody at all.” This is the most common response I received from my respondents when they described to me their listening experience of the last chorus (Example 8) of the song. True, from a musical standpoint, the last chorus can be considered to be incurring microtonal inflections. Almost no sounding pitch can fit neatly with the established (or expected) melody. Take “你縱是未明白仍夜深一人” as an example. Li performs the last two characters 一人, meaning being single or alone, as /jat5 jan21/. The two tones involve a tonal target transition of 5–1. Ideally this corresponds to a perfect fifth, major sixth, minor seventh, or a perfect octave.⁽³¹⁾ In the song it does sound like a perfect fifth, in fact, but one that is not in the F minor/A-flat major tonality. This deviation from tonal melody, however, does not make it sound like speech either, for the reason that native Cantonese speakers do not speak that way. As the musically trained said, they would describe his voice as “compressed,” “twisted,”⁽³²⁾ or “outlandish,” [古怪] as if performing with “deformed theatricality” [戲劇咁但又畸畸地形]. It is at this juncture that musical production and perception expose their differences. Recall in the first experiments that what the interviewees did was performative. That is, both the musicians and non-musicians had to make a categorial decision to showcase what they had heard: the former group mapped the uncanny onto something certain, i.e. an imaginary melody, while the latter group attempted to imitate Li’s wailing tone (of course, with individual variations). My goal in that experiment was to tease out their thought process and to see how they engaged with or broke through a speech-melody complex. Here my goal is complementary; I wish to answer the questions: how can we describe the non-speech and non-melody experience in the chorus in terms of perception? Is it something more than “outlandish”? I suggest that we can mitigate this interpretive unease by treating it as an “anamorphic listening,” a concept that I will explore further in “Verse II.”

Desperate Screaming: The Ultimate Collapse of Speech and Melody

[3.9] “The Encyclopedia” describes the end of the song as “desperate screaming in Carriage Lau’s manner” [「劉馬車」式崩潰喪叫]. Lau is notorious for his YouTube videos in which he claimed that he was a member of the Triad, a society involving in organized crime, and that he threatened another person. In one of his banned videos, he even tried to use a knife to hurt himself. The description in “The Encyclopedia” equates Li’s screams with Lau’s madness. Li’s craziness notwithstanding, the screams signify the ultimate collapse of speech and melody. They are an outpouring of accumulated anger, disappointment, and despair. The screams, in my view, make us forget that we are still in a song; in fact, they make us question if we have, in retrospect, ever been

in one. Forgetting is important here. An interviewee said he could not help but regard everything in this song as “melody,” because that was, he believed, the singer’s intention. His resistance to go into the song is not surprising; after all, it *is* a song. His experience is not to be dismissed (and I believe many would share his view).⁽³³⁾ In fact, it holds special value in explaining why, as I mentioned, speech-melody complex, at least in Cantopop, entails one’s ecological hermeneutic engagement with sound and categories. As such, when a respondent said, “I have completely no idea what is going on at the end of the song,” [我完全唔知最後嘅做緊乜] I replied, “Great! That means we are in need of an idea—”

Verse II: Anamorphic Listening

[4.1] The idea, or more precisely, the concept, is anamorphosis, a term that describes a distorted, incommensurate perspective which demands the spectator’s submission to distortion itself in order to reconstitute the perceived object—thus the “return” (*ana-*) to the “form” (*morphe*). The earliest examples of the visual technique can be traced back to Leonardo da Vinci’s drawings in the *Codex Atlanticus* (1483–1518), but Hans Holbein’s “The Ambassador” (1533) is, generally speaking, the most widely known example (Strachan [1969] 1976). In **Example 9**, what strikes the viewer’s eyes at first glance might be the two French ambassadors or their gaze. Immediately the contrast between the uncanny object at the bottom and the rest of the painting grabs one’s attention. Everything other than the object floating freely on the floor seems to be so perfectly proportioned and portrayed that one might question whether the object at the bottom is real. One can also make a heuristic guess that it might be a distorted skull. In fact, it is considered to be anamorphic art because only when one sees “The Ambassador” from a certain vantage point, in this case, from high on the right, can one perceive the skull, as shown in **Example 10**. One needs therefore to shift between perspectives when appreciating the painting. This shifting of perspectives, as Stephen Greenblatt puts it, is a “back-and-forth movement, a constant forming and reforming” (Greenblatt 1980, 23). For Daniel C. Sherer the painting is more than the visual labor entailed; it is, to a viewer, an enigma. “This enigma,” he writes, “prompts us to acknowledge an implicit tension between part and whole in the visual sign, an internal disequilibrium between the picture’s minute registration of the real and its equally meticulous disarticulation of reference” (Sherer 2000, 104). Perspective, in his view, is not only a visual projection, but a *hermeneutic* engagement, an “art of secrecy” that invites the spectator to ruminate on the *process* of seeing.⁽³⁴⁾

[4.2] Anamorphosis has influenced theories of perspective in architecture (e.g., Athanasius Kircher),⁽³⁵⁾ psychoanalysis (e.g., Jacques Lacan),⁽³⁶⁾ philosophy (e.g., Gottfried Wilhelm von Leibniz),⁽³⁷⁾ critical theory (e.g., Slavoj Žižek),⁽³⁸⁾ and computer interfaces,⁽³⁹⁾ but not music studies. This should not surprise music scholars, for the technique is borne out of silent contexts: the play with shadow projections, the foreshortening of perspectival planes, and the creation of a synthetic space that prizes illusion over reality. After all, anamorphosis is concerned with the visual, not the aural.⁽⁴⁰⁾ But it is the *concept*, not the visual technique of anamorphosis that I am channeling here. There is a critical common ground that seeing and hearing shares: time. This does not only apply to the process of perception (seeing or hearing), but also to the process of interpretation. That is, the process of hearing and seeing *as*.⁽⁴¹⁾ It is therefore the temporal, processual experience of anamorphosis that makes it a malleable concept diffused in different disciplines and fields, and I suggest that it can be a valuable tool that music scholars can play with, and therefore test, challenge, and modify. In what follows I illustrate three interlacing aspects of what I call “anamorphic listening” to enrich our unfinished discussion on the chorus of “Silly Woman.”

[4.3] (1) Hearing-as: In listening to the chorus of “Silly Woman,” respondents experienced the shifting illusion of speech and melody as if they were listening to a “distorted” sound object that is neither speech-like nor melody-like, but that is also both speech-like and melody-like. They thereby entered a space of liminality such that the perceived sound becomes impossible to define. In the first experiment, what some did was to direct their energies to creating a particular sonic category (for instance, speech). This is not to say that they favored one at the expense of the other; rather, this is to say that they valued different things in the process of imagination (see paragraph [3.5]).

The first experiment exposes a crucial difference between visual anamorphosis and sonic anamorphosis. In the former, as in the case of “The Ambassador,” one “returns” (*ana-*) to a “normal” skull if one looks at it from a particular perspective. But the word “distorted” is at best rhetorical, at worst inappropriate, in describing music that is both speech-like and melody-like, and neither speech-like nor melody-like. For in the sonic realm, there is no “point” for one to hear, nor is there a “normal” sound that one would expect. Anamorphic listening, I suggest, is not about at which point one listens to some “distorted” sonic object, but about the processual experience of negotiating the hermeneutic values of two *hearing-ases*.

[4.4] The first hearing-as is the hearing of sound as a speech-melody complex, in which one prizes the process of categorial molding as hermeneutically more important than other categorial truths. To put it analogically (but crudely), seeing “The Ambassador” as a process of perspectival distortion and recalibration is similar to listening to the “sound” in “Silly Woman” as a “speech-melody complex.” The latter is not concerned with a singular idea but with a complex ecology of relations between tones and tunes, music and language, illusion and truth, hearing and listening that one experiences and also imagines. This ecology of relations is then reflected back into the music itself, such that listening becomes an object of analysis in which the subject is inescapably implicated. The second hearing-as is the hearing of sound as separation of categories—speech or melody (not unlike seeing the skull as “normal”). This is when one grants more hermeneutic value to categorial distinction than to the process of categorial molding. Deciding, for example, whether the first line of “Silly Woman” is speech or melody is hermeneutically more important in this hearing-as, perhaps in that categorial lucidity is already an *a priori* assumption in listening to a song, as one of my interviewees maintained, or that one’s hermeneutic engagement with the process of categorial molding may detract from one’s pleasure of listening. Anamorphic listening is the negotiation of hermeneutic values between the two hearing-ases, much like the shift between perspectives in “The Ambassador.” It makes listening self-reflexive, creates “an internal disequilibrium” in one’s auditory experience, and renders objective realities one’s participatory event in one’s listening and hermeneutic valuing.

[4.5] (2) Creation and Imagination: We may recall that the musicians described Li’s voice as “outlandish.” The sonic object in question (the chorus), to my respondents’ ears, was enigmatic; the listeners were unsure as to how to define the object, to the extent that it demanded interpretations. These interpretations must create something new, that is, they must transform the inexplicable into the comprehensible by imagination. In our first experiment, one imagined speech to be melody and vice versa. As such, the return to the “normal” form is not the return to the same, but to an illusion of the same. Or, to use Christopher Hasty’s words, “‘the same’ is sustained by difference,” and the new “[comes] out of” (or emerges from) and “against the old” (Hasty 2019, 27–8).

[4.6] (3) Non-assertion: Anamorphic listening entails an apophatic learning process that does not assert, but questions or even negates.⁽⁴²⁾ As Greenblatt writes, the effect of anamorphosis is “to question the very concept of locatable reality upon which we conventionally rely in our mappings of the world, to subordinate the sign systems we so confidently use to a larger doubt” (Greenblatt 1980, 20–21). In our case, anamorphic listening is concerned with empathizing with the sonic worlds of speech and melody, hearing them as relational semiotic containers emerging from or breaking through a speech-melody complex. This is why we cannot speak of “perspective” in anamorphic listening as in visual anamorphosis. “The word ‘perspective,’” as Eugene Gendlin reminds us, “assumes that the environment is something merely viewed, not interacted with and behaved in. First person process is not a perspective” (2009, 349). Listening to the chorus of Li’s song is a radical questioning of pure illusionism—one’s very listening—that in turn becomes materialized in one’s lived experiences.

Chorus II: FAMA’s “No Boundaries in the Sea of Knowledge” (2007)

[5.1] The song “No Boundaries in the Sea of Knowledge,” performed by the two-man group FAMA, does not propagate the idea, as the title may otherwise imply, that since there are no limits to knowledge, one has to pursue it endlessly. Rather, it satirizes the ossified educational system in Hong Kong in which students only bury themselves into books. Apart from showing empathy for

Hong Kong students' plight, ultimately, the song conveys the Chinese counterpart of the saying, "One that travels far knows much" [讀萬卷書不如行萬里路].

Language Play in Cantopop

[5.2] The song begins with the following verse (**Video Example 2**). I have notated the dictionary speech tones as follows (0:35–0:55):

農夫書院 美好人生
nung21 fu55 syu55 jyun22 mei23 hou25 jan21 sang55
有兩個訓導陸永與C君
jau23 loeng23 go33 fan33 dou22 luk2 wing23 jyu23 C gwan55
我們都要努力做人
ngo23 mun21 dou55 jiu33 nou23 lik2 zou22 jan21
個個會考三十分
go33 go33 wui22 haau25 saam55 sap2 fan55

Translation: "In FAMA college, everyone leads a happy life. The two disciplinary masters are 6wing and C Kwan. We have to work hard in life and be high-flyers in public exams."

[5.3] But FAMA performs them as:

農副書怨 眉好因搯
nung21 fu33 syu55 jyun33 mei21 hou33 jan55 sang33
有兩個訓導陸永遇C棍
jau23 loeng23 go33 fan33 dou21 luk2 wing23 jyu22 C gwan55
鵝滿都要努力做人
ngo21 mun23 dou55 jiu33 nou23 lik2 zou22 jan21
個個會孝三十分
go33 go33 wui22 haau33 saam55 sap2 fan55

[5.4] I am not interested here in looking at the changed tones and their corresponding semantics (in fact, some respondents said they completely understood the lyrics),⁽⁴³⁾ but rather on the responses I received as to the divergent opinions of the song's speech-melody relationships. One musician said the beginning of the *song* reminded her of the language play in colloquial Cantonese. She said, "We often play with our language [Cantonese] by exaggerating speech tones to imitate Mandarin or English."⁽⁴⁴⁾ Her statement underlines that language hybridity in Cantopop does not involve merely the inclusion or insertion of English and/or Mandarin; it is also manifest by changing the tones *in* Cantonese. For instance, in changing 我們 /ngo23 mun21/ ("we") to 鵝滿 /ngo21 mun23/ (nonsense; literally, "goose full"), FAMA imitates the Mandarin 我們 /wǒ mén/. Accordingly, one might claim that at that juncture, a cross-contextuality of languages is at play in a speech-melody complex. FAMA sings in Cantonese, but speaks in Mandarin *at the same time*.⁽⁴⁵⁾

[5.5] I then asked her a follow-up question, "You said the opening lines in the song remind you of language play in our colloquial speech. Would you mind clarifying that? Do you consider it as speech within a song?" She faltered, claiming that it might be contradictory to say that. She then said, "I meant the melody . . . no, the song as a whole . . . Wait. Right, if you see it as language play then it may be considered speech. But compared to the rap in the chorus, it seems to be more melodic." What captivated me during the interview was her response with regard to the contradictory speech-melody complex vis-à-vis her immediate shift to a more fluctuating stance. We can perhaps look at what she had thought speech and melody are in Example 4. She is the musician (number 3) who described speech as "semantically centered" and melody as "speech with musical intervals." There is nothing about language play in her first explanations. The construction process of speech/melody as categories, as demonstrated from her hesitation and her second attempt to explain them, is always shifting. It involved comparisons, perspectival thinking, and linguistic contextualization. It also shows that the larger the variety of information to which one is alert, the less important is the question of what is speech and what is melody. Speech and melody, as part of the same conceptual package of sounding tunes, form a complex in which

listeners take into account different contextual factors to decide which is to be foregrounded. The shift in this interviewee's answer from melody to speech exemplifies this foreground/background conceptual (and perhaps also perceptual) motion. This also explains why some respondents held a completely different view from hers. One said the verse is "an utter distortion of Cantonese speech tones by the musical melody;"⁽⁴⁶⁾ another said that it is "muddled but at the same time clear."⁽⁴⁷⁾ These respondents did not take into account the language play, but turned their attention to semantics, notably the time needed to process the meaning of the lyrics. What gels together their views is the play of languages, sounds, and listening in a speech-melody complex.⁽⁴⁸⁾

Listening Into, Listening Out Of

[5.6] Our analysis of Li's "Silly Woman" is focused on particular moments, and especially on words. It requires a kind of listening *into*, a highly attentive listening. But is a speech-melody complex only a matter of attending to specific moments in Cantopop? How useful is it? What about listening *out of*? I attempt to answer these questions by analyzing the following line in FAMA's song, which is located after the rap section (1:50–1:53):

放得暑假我個人鬆晒

fong33 dak5 syu25 gaa33 ngo23 go33 jan21 sung55 saai33

Translation: "Summer holidays are coming soon. I can relax."

[5.7] This line forges a gray area which both speech and melody inhabit on two levels: "listening into" and "listening out of." We may listen into the interaction between tonal target transitions and melodic intervals, which is, as I mentioned in my analysis of "Silly Woman," one of the factors contributing to a speech-melody enigma; or we may listen out of the sonic qualities and effects of the tones and listen into the way the lyrics are written. It is a written form of Cantonese; it is a colloquial form of writing that equates speaking and writing. This familiarity, or even intimacy, to Cantonese speakers is a kind of hailing from the performers: a calling out not to listen to the tones, but to listen to them speak—or, to listen *into* the language. In this way, listening into and out of are not situated in a dichotomy. When we listen out of this (say, tones), we listen into that (say, language), and vice versa. A speech-melody complex embraces these in-and-out, needling listening processes.

[5.8] Another layer of complication is rhythm (which deserves its own study). For some the line is "melodic" only because the way each character was accorded an equal interval of time sounded strange to them. A melody, for them, imposes rhythmic constraints to colloquial speech. I asked the interviewees to re-produce this sentence, and to compare their delivery with that in the song. Many were surprised (that this was a common emotional response reveals that they did not think about how they usually talk when listening to the song). They found their production and FAMA's performance to be identical, thereby relating the performance to colloquial speech; others were troubled by the rhythmic regularity of the original performance, though they did not therefore claim that it was a melody. It sounded to them like "unnatural" speech. I shared with both camps the opposite side's views and asked what they thought the supporting reasons behind the other camp's claims were. They understood why other Cantonese speakers would think otherwise. For example, some respondents said there is a slight possibility that they would divide the words evenly when they speak. Others claimed that what "colloquial speech" sounds like differs from speaker to speaker.

[5.9] One respondent, who is not a classically trained musician, claimed that he "would not listen to rhythm" [我唔聽rhythm㗎] and that therefore the line "放得暑假我個人鬆晒" sounded like colloquial speech to him. I could not help but raise my eyebrows, in part, thinking in retrospect, because I carried with me some music-theoretical premises as to what "rhythm" might be, and in part because the causality between "not listening to rhythm" and "colloquial speech" piqued my curiosity. When I asked for clarification, he explained:

I'm not saying that rhythm is unimportant, but you won't think about beats all the time when you listen to music. You don't have that many pairs of ears to attend to so

many things. I might pay attention to rhythm if it's special or something, but in general I tend to lyrics and melody [歌] more. When I listen to the line “放得暑假我個人鬆晒” I don't quite think of rhythm, because it's just colloquial speech.⁽⁴⁹⁾

[5.10] This respondent's repeated emphasis on the connection between “rhythm” and “colloquial speech” revealed two crucial differences between his listening process and that of others. First, while “rhythm” served as a factor influencing the others' negotiation of speech-melody relationships, for him, the category of speech obviated the need to attend to “rhythm” in music. Second, in his view, listening to “rhythm” puts a burden on the perceptual load in listening to Cantopop songs insofar as lyrics and melody have already occupied much of his attention capacity (and he does not have so many pairs of ears). It is only when the “rhythm” is “special” that he would “listen” to it.

[5.11] I asked if he was trying to refer to the pulse of a musical passage (i.e., the beat), and he answered in the affirmative. I then wondered why the beat in “放得暑假我個人鬆晒” is not “special” if he considered it to be colloquial speech. I showed him how the downbeats fall onto the verb particle “得” and the classifier “個,” which are grammatical particles that native Cantonese speakers would not usually stress. The downbeat also falls onto the character “假”; when speaking the word compound “暑假” (which means “summer holidays”) native Cantonese speakers usually either stress the first character rather than the second, or grant equal weight to both characters. He then exclaimed, “I couldn't hear it [before]; now it doesn't sound like colloquial speech!”⁽⁵⁰⁾ Our conversation (which then evolved into the two-way street between music theory and listening) points to the potential relations between the needling listening process and anamorphic listening: it is not “beat” *per se*, but the process of listening into and out of which, that makes the line a speech-melody enigma, and such a process will vary in every listen.

Outro

[6.1] In *When Words Sing* (1970), Murray Schafer reported the following anecdote about a seven-year-old girl Miranda's made-up story and offered his commentary:

“Once there was a little old man. He wanted to get a piece of volcano rock. So he got all his hiking equipment ready. Then he hiked up the mountain. It took him twenty years. When he got to the top he took off the lid. Oooooahooooo! Inside there was a ghost and he moaned and he groaned. The man ran away as fast as he could and he said, “I'm never going to climb volcanoes again!”

Miranda recites her unusual story with intense emotion. Words like ‘moaned’ and ‘groaned’ are so highly inflected and attenuated they are almost chanted. The ‘Oooooahooooo’ is a pure glissando melody. Miranda knows that words are magic invocations and can cast spells. So she exorcizes them with music. Of course her teachers will correct all this in another year or two by muzzling her to the printed page (Cited in [Leeuwen 1999](#), 3–4).

[6.2] In Western art history, there has been a long tradition of story-telling and poetry recitation in which one hears melody in speech, and another tradition of melodrama and recitative in which one hears speech in melody. In this article, however, I, together with my interviewees, tell a story that synthesizes speech and melody into an “unseparated multiplicity,” ([Gendlin 2018](#), 35, 44) a speech-melody complex that embraces the forming of the two recalcitrant and mischievous sounds. These do not resolve hastily into categories, but form a complex that allows Cantonese users to wander in, contemplate upon, experience, and question. This is not to say that the categorial boundary between speech and melody cannot exist, but it exists only to be crossed, to be unmoored. Categories of speech and melody, as such, merely function as temporary explications of auditory experience. It is when speech and melody are separated out from one another that they gain their own identity—or really, exist to *someone*; it is also when one asserts some sonic object to be speech or melody that it *becomes* speech or melody. In other words, a speech-melody complex hoards speech's and melody's potentialities of interactions, integration, divergence, and indeed, messiness.

It follows, then, that describing sonic patterns as speech or melody is more than an assertion: it is a breaking through of a speech-melody complex and a creation of concepts through separation. A speech-melody complex does not correspond to an external reality out there, but is constantly interacting with the environment, renewing itself, and leading us to places that few have travelled.

[6.3] I also tell a meta-story of anamorphic listening that captures the negotiation of the hermeneutic values of two hearing-ases in Cantopop: the listening to Cantopop as a speech-melody complex, and as a separation of categories. Cantopop songs, I think, present native Cantonese speakers with a speech-melody enigma (from the Greek *ainos*: story, fable) that invites us to submit ourselves to the hermeneutic engagement of sounds, the repetition (*ana-*) of the new (*morphe*), and the (dis)articulation of habitual speech-melody relationships. Every Cantopop song, as I hear it, is a musical-linguistic artwork.

[6.4] Of course, in this article, I can only scratch the surface of a rich, yet under-discussed, genre; there is so much to explore in Cantopop to theorize about the relationships between music and language. We may, for example, further ask: How might the concepts introduced in this article influence the process of setting text to melody in Cantopop, and vice versa? What is the role of affect and voice in speech-melody interactions in Cantopop? Going beyond Cantopop, we may ask: how does a speech-melody complex work in other art genres in Cantonese, such as Cantonese opera, Cantonese rap, spoken word poetry, and contemporary music? Do we need some other concepts for them?

Edwin K. C. Li
Harvard University
Department of Music
3 Oxford St.
Cambridge, MA 02138
edwin_li@g.harvard.edu

Works Cited

- Agawu, Kofi V. 1988. "Tone and Tune: The Evidence for Northern Ewe Music." *Africa: Journal of the International African Institute* 58 (2): 127–46.
- Baart, Joan L. G. 2004. "Tone and Song in Kalam Kohistani (Pakistan)." In *On Speech and Language: Studies for Sieb G. Nootboom*, edited by Hugo Quené and Vincent van Heuven, 5–16. Netherlands Graduate School in Linguistics.
- Barry, Johanna G., and Peter J. Blamey. 2004. "The Acoustic Analysis of Tone Differentiation as a Means for Assessing Tone Production in Speakers of Cantonese." *The Journal of the Acoustical Society of America* 116 (3): 1739–48.
- Bauer, Robert, and Paul Benedict. 1997. *Modern Cantonese Phonology*. Mouton de Gruyter.
- Carter-Ényì, Aaron, and Quintina Carter-Ényì. 2020. "Melodic Language and Linguistic Melodies," *SMT-V* 6, no. 5.
- Chai, David. 2017. "Zhuangzi and Musical Apophasis." *Dao: A Journal of Comparative Philosophy* 16 (3): 355–70.
- Chan, Marjorie K. M. 1987. "Tone and Melody in Cantonese." In *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 26–37.
- Cheung, Kwan-hin. 2016. *Cong Zhonggu Sisheng dao Xiandai Yueyu Liudiao* 從中古四聲到現代粵語六調 [From the Ancient Four Tones to the Six Tones in Modern Cantonese]. New Horizons in the Study of Chinese: Dialectology, Grammar, and Philology, 361–84.
http://www.cuhk.edu.hk/ics/clrc/english/pub_yue.html.

- Chik, Alice. 2010. "Creative Multilingualism in Hong Kong Popular Music." *World Englishes* 29 (4): 508–22.
- Chow, Man-ying. 2012. "Singing the Right Tones of the Words: The Principles and Poetics of Tone-melody Mapping in Cantopop." M. Phil. thesis, The University of Hong Kong.
- Chu, Yiu-wai. 2017. *Hong Kong Cantopop: A Concise History*. Hong Kong University Press.
- Day, William, and Victor J. Krebs, eds. 2010. *Seeing Wittgenstein Anew*. Cambridge University Press.
- Deutsch, Diana. 2019. *Musical Illusions and Phantom Words: How Music and Speech Unlock Mysteries of the Brain*. Oxford University Press.
- Deutsch, Diana, et al. 2011. "Illusory Transformation from Speech to Song." *The Journal of the Acoustical Society of America* 129 (4): 2245–52.
- Dubiel, Joseph. 2017. "Music Analysis and Kinds of Hearing-As." *Music Theory and Analysis (MTA)* 4 (2): 233–42.
- Feld, Steven. 1974. "Linguistic Models in Ethnomusicology." *Ethnomusicology* 18(2): 197–217.
- . 1984. "Communication, Music, and Speech about Music." *Yearbook for Traditional Music* 16: 1–18.
- Fok Chan, Yuen-yuen. 1974. *A Perceptual Study of Tones in Cantonese*. University of Hong Kong.
- Fox, Anthony, Kang-Kwong Luke, and Owen Nancarrow. 2008. "Aspects of Intonation in Cantonese." *Journal of Chinese Linguistics* 36 (2): 321–67.
- Gabora, Liane, Eleanor Rosch, and Diederik Aerts. 2008. "Toward an Ecological Theory of Concepts." *Ecological Psychology* 20 (1): 84–116.
- Gendlin, Eugene. 2009. "What First and Third Person Processes Really Are." *Journal of Consciousness Studies* 16 (10–11): 332–62.
- . 2018. *A Process Model*. Northwestern University Press.
- Goodman, Nelson. 1978. "When is Art?" In *Ways of Worldmaking*, 57–70. Hackett Publishing Company.
- Greenblatt, Stephen. 1980. *Renaissance Self-Fashioning: From More to Shakespeare*. University of Chicago Press.
- Harrell, Stevan. 1979. "The Concept of Soul in Chinese Folk Religion." *Journal of Asian Studies* 38 (3): 519–28.
- Hasty, Christopher. 2019. "Thinking with and About Rhythm." In *Thought and Play in Musical Rhythm: Asian, African, and Euro-American Perspectives*, edited by Richard K. Wolf, Stephen Blum, and Christopher Hasty, 20–54. Oxford University Press.
- Ho, Wing-see Vincie. 2010. "A Phonological Study of the Tone-melody Correspondence in Cantonese Pop Music." Ph.D. diss., The University of Hong Kong.
- Huang, Zhi-hua. 2003. *Yueyu Geci Chuangzuo Tan 粵語歌詞創作談 [On Writing Cantonese Lyrics]*. Joint Publishing (H.K.) Co. Ltd.
- . 2020. *Wenzi Shenglü yu Yueyuge Chuangzuo 文字聲律與粵語歌創作 [Tone, Lü, and Creative Processes of Cantopop Songs]*. 懿津出版企劃有限公司.
- Ketkaew, Chawadon, and Pittayawat Pittayaporn. 2014. "Mapping between Lexical Tones and Musical Notes in Thai Pop Songs." In *28th Pacific Asia Conference on Language, Information and Computational pages*, 160–69.

- Khouw, Edward, and Valter Ciocca. 2007. "Perceptual Correlates of Cantonese Tones." *Journal of Phonetics* 35 (1): 104–17.
- Kirby, James, and D. Robert Ladd. 2016. "Tone-melody Correspondence in Vietnamese Popular Song." In *Tonal Aspects of Language: 5th International Symposium*, 48–51.
- Lacan, Jacques. 1981. *The Four Fundamental Concepts of Psychoanalysis*, edited by Jacques-Alain Miller, translated by Alan Sheridan. W.W. Norton & Company.
- . 1992. *The Ethics of Psychoanalysis 1959–1960: The Seminar of Jacques Lacan (Book VII)*, edited by Jacques Alain-Miller, translated by Dennis Potter. W.W. Norton & Company.
- Leeuwen, Theo van. 1999. *Speech, Music, Sound*. Macmillan Press Ltd.
- Leung, Wai-sum. 2016. "Hong Kong Cantonese in Cantopop from Phonological and Lexical Perspectives." M.A. thesis, The University of Hong Kong.
- Li, Xin-kui, Wang Jia-jiao, Shi Qi-sheng, Mai Yun, and Chen Ding-fang. 1995. *Guangzhou Fangyan Yanjiu 廣州方言研究 [A Study of the Guangzhou Dialect]*. Guangdong People's Publishing House.
- Liu, An. 2010. *The Huainanzi: A Guide to the Theory and Practice of Government in Early Han China*, translated and edited by John S. Major, Sarah A. Queen, Andrew Seth Meyer, and Harold D. Roth, with addition contributions by Michael Puett and Judson Murray. Columbia University Press.
- Matthews, Stephen, and Virginia Yip. 2011. *Cantonese: A Comprehensive Grammar 廣東話語法*, 2nd edition. Routledge.
- McPherson, Laura, and Kevin M. Ryan. 2018. "Tone-tune Association in Tommo So (Dogon) Folk Songs." *Language* 94 (1): 119–56.
- Ngai, Sianne. 2012. *Our Aesthetic Categories: Zany, Cute, Interesting*. Harvard University Press.
- Page, Christopher. 1993. "Johannes de Grocheio, the *Litterati*, and Verbal *Subtilitas* in the *Ars Antiqua Motet*." In *Discarding Images: Reflections on Music and Culture in Medieval France*, 65–111. Clarendon Press.
- Patel, Aniruddh D. 2008. *Music, Language, and the Brain*. Oxford University Press.
- Pérez-Gómez, Alberto, and Louise Pelletier. 1996. *Anamorphosis: An Annotated Bibliography*. McGill-Queen's University Press.
- Powers, Harold S. 1980. "Language Models and Musical Analysis." *Ethnomusicology* 24 (1): 1–60.
- Qian, Yao, Tan Lee, and Frank K. Soong. 2007. "Tone Recognition in Continuous Cantonese Speech Using Supratone Models." *The Journal of the Acoustical Society of America* 121 (5): 2936–45.
- Ravnik, Robert, et al. 2014. "Dynamic Anamorphosis as a Special, Computer-Generated User Interface." *Interacting With Computers* 26 (1): 46–62.
- Rogalsky, Corianne, et al. 2011. "Functional Anatomy of Language and Music Perception: Temporal and Structural Factors Investigated Using Functional Magnetic Resonance Imaging." *The Journal of Neuroscience* 31 (10): 3843–52.
- Rosch, Eleanor. 1999. "Reclaiming Concepts." *Journal of Consciousness Studies* 6 (11–12): 61–77.
- Schneider, Marius. 1961. "Tone and Tune in West African Music." *Ethnomusicology* 5 (3): 204–15.
- Sherer, Daniel C. 2000. "Anamorphosis and the Hermeneutics of Perspective from Leonardo to Holbein, 1490–1533." Ph.D. diss., Harvard University.
- Strachan, W. J. 1976. *Anamorphic Art*. Chadwyck-Healey Ltd. Translation of Baltrušaitis, Jurgis. 1969. *Anamorphoses ou magie artificielle des effets merveilleux*. Oliver Perrin éditeur.

- Tanese-ito, Yoko. 1988. "The Relationship between Speech-tones and Vocal Melody in Thai Court Song." *Musica Asiatica* 5: 109–39.
- The Encyclopedia of Virtual Communities in Hong Kong. "李逸朗." <http://evchk.wikia.com/wiki/李逸朗>. Last accessed on 29 November 2019.
- The Hong Kong Polytechnic University. 2004. "Standardization of Cantonese Romanization and the Building Blocks for Phonetic Based Applications" [粵語漢字羅馬拼音標準與語音應用的基礎模塊]. <http://www.iso10646hk.net/jp/index.jsp>. Last accessed on 29 November 2019.
- Tierney, Adam, et al. 2013. "Speech versus Song: Multiple Pitch-Sensitive Areas Revealed by a Naturally Occurring Musical Illusion." *Cerebral Cortex* 23 (2): 249–54.
- Tierney, Adam, Aniruddh Patel, and Mara Breen. 2018. "Acoustic Foundations of the Speech-to-Song Illusion." *Journal of Experimental Psychology: General* 147 (6): 888–904.
- Wittgenstein, Ludwig. 1953. *Philosophical Investigations*. Translated by G. E. M. Anscombe. Blackwell.
- Wong, James. 1997. *Yueyu Liuxingqu de Geci Chuangzuo* 粵語流行曲的歌詞創作 [Writing of Verses for Cantonese Popular Songs], Occasional Paper Series no. 2. Lingnan College Centre for Literature and Translation.
- Wong, Patrick C. M. 1999. "The Effect of Downdrift in the Production and Perception of Cantonese Level Tone." In *Proceedings of the XIVth International Congress of Phonetic Sciences*, 2395–98.
- Wong, Shek-ling. 1941. *A Chinese Syllabary Pronounced According to the Dialect of Cantonese*. Lingnam University. <http://humanum.arts.cuhk.edu.hk/Lexis/Canton>.
- Yiu, Suki S. Y. 2013. "Cantonese Tones and Musical Intervals." In *Proceedings of the International Conference on Phonetics of the Languages in China, ICPLC-2013*, 155–58.
- Yu, Mengxia, et al. 2017. "The Shared Neural Basis of Music and Language." *Neuroscience* 357: 208–19.
- Yung, Bell. 1989. *Cantonese Opera: Performance as Creative Process*. Cambridge University Press.
- Zhang, Ling. 2016. *Intonation Effects on Cantonese Lexical Tones in Speaking and Singing*. LINCOM Studies in Phonetics.
- Zhu, Xiao-nong, and Wang Cai-yu, eds. 2015. *The Oxford Handbook of Chinese Linguistics*. Oxford University Press.
- Žižek, Slavoj. 2000. "Melancholy and the Act." *Critical Inquiry* 26 (4): 657–81.
- 粵音資料集叢 <https://jyut.net>. Last accessed on 29 November 2019.

Footnotes

* An earlier version of this article was presented at the New England Chapter for the American Musicological Society (Spring 2019) at the College of the Holy Cross. I am grateful to the anonymous interviewees who devoted their time to bringing this project into fruition, Richard Wolf who saw the embryo of this article, Alex Rehding who read the script enthusiastically, Christopher Hasty who shaped my thinking in many ways, Samuel Chan who saw potentialities in my project, and the two anonymous reviewers of this journal who offered invaluable suggestions. [Return to text](#)

1. Recently, Diana Deutsch has published a monograph on this topic, summoning both historical and scientific perspectives in the investigation of the relationship between music, speech, and the brain. See [Deutsch 2019](#). [Return to text](#)

2. However, as Adam Tierney and others point out, not all acoustical stimuli intended to be heard as speech are perceived as melody. In fact, certain cues in acoustical stimuli might prompt one to hear speech as melody: the presence of flat pitches, syllable pitches that bear resemblance to musical scale structure, rhythmic regularity, repetition, and more. See [Tierney, et al. 2013](#); [Tierney, Patel, and Breen 2018](#).

[Return to text](#)

3. Yu Mengxia et al. [2017](#) shows that semantic processing of language and melodic analysis are positively correlated with spontaneous activities in various parts of the brain. For a counterargument, see [Rogalsky, et al. 2011](#).

[Return to text](#)

4. See [Deutsch 2019](#), Chapter 10.

[Return to text](#)

5. In this article I do not distinguish concepts from categories, for, as Eleanor Rosch has it, they are both non-presentational but a participatory part of the mind that only exists in complex webs of relations; see [Rosch 1999](#).

[Return to text](#)

6. See, for example, [Powers 1980](#); [Feld 1974, 1984](#).

[Return to text](#)

7. Cantopop has assumed different meanings in time. This article takes it to refer to songs that are considered by native Cantonese speakers to be Cantopop. For a historical study of the term, see [Chu 2017](#).

[Return to text](#)

8. In this article “listeners” refers specifically to native Cantonese speakers who also understand and speak Mandarin and English. As I shall show in “Chorus II,” to appreciate the complexity of the issue of speech-melody in Cantopop, native Cantonese speakers are expected to be aware of the language play and the practice of code-switching between the three languages.

[Return to text](#)

9. See [Matthews and Yip 2011](#), 28. This distinction is not without problems, for isolated pronunciation of speech tones is not tantamount to context-free tone production. For example, I can pronounce a tone in isolation, but also in contexts such as speaking to a friend, narrating a story, or asking a question. One can claim, then, that basic tones exist in theory but not in practice.

[Return to text](#)

10. For more on this phonological categorization, see [Zhu and Wang 2015](#), especially the sections “History” and “Phonetics and Phonology.” In this tone-letter system, tone names and letters do not necessarily agree. For example, although there is no change in tonal contour in /33/, it is categorized as a “high falling tone”; by contrast, /21/ is a “level tone.” This should not be considered a problem, but an attempt to capture the practical usages of Cantonese in particular regions. As such, different systems assign different tone letters to the same tone name. See, for instance, [Li, et al. 1995](#) and [Cheung 2016](#). But this system was soon out-popularized by another Cantonese-specific phonological system *jyutping* 粵拼, a Cantonese Romanization system developed by the Linguistic Society of Hong Kong in 1993. It assigns each tone a number, ranging from 1 to 6, in the order from high level to mid-low level. The number, however, has no correlation whatsoever with pitch contour or relative pitch level. For the sake of illustrating the relative pitch motion of speech tones in this article, I decided to adopt the system of tone letters. For Cantonese Romanization, I consulted the online *jyutping* database developed by The Hong Kong Polytechnic University in 2004, “Standardization of Cantonese Romanization and the Building Blocks for Phonetic Based Applications” [粵語漢字羅馬拼音標準與語音應用的基楚模塊], <http://www.iso10646hk.net/jp/index.jsp>. For cross-reference I consulted the online version of Wong Shik-ling’s *A Chinese Syllabary Pronounced According to the Dialect of Cantonese*,

<http://humanum.arts.cuhk.edu.hk/Lexis/Canton/>.

[Return to text](#)

11. There is no consensus among linguists as to the number of speech tones in Cantonese. Robert Bauer and Paul Benedict (1997, 11) claim that there are six to seven tones, while Matthews and Yip (2011, 11) claim that there are six. The number of tones in Cantonese, in fact, is predicated on many factors, including the region from which the Cantonese speaker originates (e.g., Hong Kong, Guangdong, or even multiple regions), the language traits (e.g., pitch contour, pitch level, and so forth) that the system highlights, the inclusion of variants, and the grouping of tones (e.g., entering tones should, as some linguists suggest, be grouped with non-entering ones).

[Return to text](#)

12. Scholars have attempted to answer this question in musics across cultures in which tonal languages are practiced. See, for example, [Carter-Ényì and Carter-Ényì 2020](#); [Kirby and Ladd 2016](#); [Ketkaew and Pittayaporn 2014](#); [Baart 2004](#).

[Return to text](#)

13. 「現在，香港人的耳朵尖刻得多了，再不容許這種滿口倒字的粵語歌詞存世。」 Translated by and cited in [Chow 2012](#), 7, 103.

[Return to text](#)

14. 「字配上了樂音，聲調不變，就是填得對。字聲聽出來，變了另一個字，就要換字，換到聽唱準確為止。」 Translated by and cited in [Chow 2012](#), 37.

[Return to text](#)

15. For a recent theory of tone-melody mapping and compositional process in Cantopop, see [Huang 2020](#).

[Return to text](#)

16. In the same spirit, Suki S. Y. Yiu claims that Cantonese tones can be understood as musical intervals; see [Yiu 2013](#).

[Return to text](#)

17. Important perceptual studies of Cantonese include [Fok Chan 1974](#); [Barry and Blamey 2004](#); [Khouw and Ciocca 2007](#); [Qian, Lee, and Soong 2007](#). One important takeaway from these studies is that perception of a Cantonese speech tone is largely based on communicative context.

[Return to text](#)

18. This runs counter to Huang's claim that "the directionality of tonal transition has to be the same as the sung melody" (「字音高走向與旋律音階走向要一致，不可逆向。」). [Huang 2003](#), 106.

Cited in [Chow 2012](#), 43, with "致" being modified here to "致."

[Return to text](#)

19. In many cultures, speech tones and musical melody are considered separately along the line of their correspondence. Yoko Tanese-ito, for instance, states that the speech tones of the Thai language are reflected in the basic melody of one type of Thai court song but not another (1988). Marius Schneider (1961) says: "If a word is to be grammatically intelligible, the individual syllables cannot be sung arbitrarily high or low. Speech tone and musical tone must be definitely correlated" (204). A recent study that adopts a similar approach to analyzing speech-melody relationships in Tommo So (a Dogon language of Mali) folk songs is [McPherson and Ryan 2018](#).

[Return to text](#)

20. My argument regarding the ecology of concepts is inspired by [Gabora, Rosch, and Aerts 2008](#).

[Return to text](#)

21. This is inspired by Nelson Goodman's (1978) article "When is Art?"

[Return to text](#)

22. The literal translation is "The song of the big hotel." "The big hotel" refers to the funeral parlor.

[Return to text](#)

23. July 14th is the date of the Hungry Ghost Festival. In some Chinese traditions, “soul-calling” (喊驚, 收驚, or 收魂) refers to a wide variety of vocal techniques (e.g., shouting, reciting religious scriptures, singing) one employs to summon the human soul, which is believed to have left a person’s body and caused illness, because the soul is disturbed by ghosts. For a succinct history and cultural context of the concept of soul and the practice of soul-calling in China and elsewhere, see [Harrell 1979](#).

[Return to text](#)

24. The original Cantonese text reads: 「哇！我輸了喇，他唱得感情比我豐富得多。不過我覺得這個應該是《傻女》的劇場版。。。我覺得他蠻有噱頭、有gimmick，有娛樂性，我自己就幾喜歡這個版本。」 <http://evchk.wikia.com/wiki/李逸朗>.

[Return to text](#)

25. 「好多人留言給防止自殺會和醫生的電話給我。。。其實我想音樂的東西，不應該被規範在框框裡面。。。」 *Ibid.* The Samaritan Befrienders is an organization which provides counseling services for those who intend to commit suicide.

[Return to text](#)

26. However, in the lyrics, “lonely” is literally used to describe the “air.”

[Return to text](#)

27. I do not use “phonological variation” in the sense that the phonology of certain words has changed over time, but for the different phonological possibilities of the same character. For the former usage, see [Leung 2016](#).

[Return to text](#)

28. This is according to some respondents’ own experiences. For more information on Cantonese intonation, see [Fox, Luke, and Nancarrow 2008](#). The drop in pitch level in this case is not identical to “downdrift,” when a pitch level naturally falls from the beginning to the end of an utterance ([Wong 1999](#)).

[Return to text](#)

29. 「/ze23/ 有D archaic同pedantic囉，但喺我諗有D人會咁講嘅。」

[Return to text](#)

30. According to various sources, /ze23/ or /ze33/ are the dictionary pronunciations of the character “這.” See [粵音資料集叢](#) <https://jyut.net/query?q=這>.

[Return to text](#)

31. For a table of the most common melodic intervals for the twelve non-level tonal target transitions, see [Ho 2010](#), 79. [Huang \(2020\)](#) has recently proposed a slightly different table that includes level tonal target transitions (28–9), suggesting that the “most natural” [最自然] melodic interval for them would be a unison. Both Ho and Huang further complicate this “intuitive” proposal. See [Ho 2010](#), 60–6; [Huang 2020](#), 52–61.

[Return to text](#)

32. “Compressed” and “twisted” were given in English. “Outlandish” is my translation of the Cantonese.

[Return to text](#)

33. Some confessed to me that when listening to Cantopop, sometimes they do not pay attention to the lyrics at all, not to mention what the meaning of the song is. They merely relish the rushing sonority produced by the lyrics for aesthetic pleasure. This listening practice resonates with what some medieval audiences might have done when they listened to motets. As medievalist Christopher Page writes, “Experience suggests that the pleasure derived from hearing the words of a song does not necessarily rest upon a full (or even moderate) comprehension of their lexical meaning and syntactic relations, let alone of their ‘meaning’ in any broader sense of the term” (1993, 86).

[Return to text](#)

34. Sherer (2000, 49–80) explains this self-reflexive process of view in his discussion of German Renaissance artist Albrecht Dürer's engraving *Melencolia I* (1514). Dürer called his perspective a "*Kunst in heimlicher Perspectiva*" — an "art of secret perspective," which provokes the viewer to "reflect upon the process of viewing represented in the image."

[Return to text](#)

35. See Strachan [1969] 1976, 71–91.

[Return to text](#)

36. As Lacan writes, anamorphosis is "any kind of construction that is made in such a way that by means of an optical transposition a certain form that wasn't visible at first sight transforms itself into a readable image." Lacan 1992, 135. Elsewhere he considers anamorphosis to be the "gaze imagined by me in the field of the Other." Lacan 1981, 84.

[Return to text](#)

37. See Pérez-Gómez and Pelletier 1996, 90.

[Return to text](#)

38. See, for example, Žižek 2000. His argument that "anamorphosis undermines the distinction between objective reality and its distorted subjective perception; in it, the subjective distortion is reflected back into the perceived object itself, and, in this precise sense, the gaze itself acquires a supposedly objective existence" (659) echoes the breaking down of the subject-object binary in my discussion of anamorphic listening.

[Return to text](#)

39. See Ravnik, et al. 2014.

[Return to text](#)

40. For an exponent of "musical perspective," see Leeuwen 1999, especially Chapter 2.

[Return to text](#)

41. The idea of "hear-as" is inspired by Ludwig Wittgenstein's discussion of the difference between "seeing-that" and "seeing-as" in his *Philosophical Investigations* (1953). Put crudely, by "seeing-as" Wittgenstein refers to the activity which takes place when one is concerning oneself with the object rather than merely reporting what is there in "seeing-that." For a critical study of these ideas, see Day and Krebs 2010; for Wittgenstein's writing on them, see in particular Section 11 in Part II of the *Investigations*. For an example for the use of "hearing-as" in Western music analysis, see Dubiel (2017).

[Return to text](#)

42. For an idea of musical apophysis, see Chai 2017.

[Return to text](#)

43. This may conjure up the non-correspondence between speech tones and melody in the Cantonese song "Spring Field Kindergarten School Song;" see Chow 2012, 105–07.

[Return to text](#)

44. I would add that "imitate" might be a qualified expression. Sometimes Hongkongers exaggerate the Canto-inflected Mandarin to play with the differences between the two languages.

[Return to text](#)

45. For multilingualism and popular music in Hong Kong, see Chik 2010.

[Return to text](#)

46. [D音完全比個melody整錯晒。]

[Return to text](#)

47. [唔知up乜但又明。]

[Return to text](#)

48. Another common point of interest among my respondents was the opening lines “農夫書院 美好人生 有兩個訓導陸永與C君” and their English counterparts. When the lyrics are reiterated in English later as “FAMA college beautiful life; we have 6wing and C Kwan be your guide,” some, who first considered the opening lines to be melody, expressed later that the English version is melody, while the Cantonese one is speech. This shows how speech-melody complex is ever-changing against different contexts. Even in the progress of a song, one can retroactively reconfigure their figure-ground conceptualization. This dynamic process concerns how much contextual information one takes into consideration in shaping and parsing out the speech-melody complex. The contextual information, with which one may or may decide to foreground speech or melody, then becomes one’s conceptual inventory. When this inventory accumulates more exposure and experience, it in turn re-packages our conceptual thoughts.

[Return to text](#)

49. 「我嘅意思唔係話rhythm唔重要，但你聽歌唔會成日都諗住個beat嘅點嘍嘛，邊有咁多隻耳聽咁多嘢，會留意rhythm可能佢好特別或者唔知點，不過in general我都係留意D詞同歌多D。果句『放得暑假我個人鬆晒』我有諗咩rhythm啲，因為佢似講野多D囉。」

[Return to text](#)

50. 「我完全冇聽到，而家又好似唔似講野啦啲！」

[Return to text](#)

Copyright Statement

Copyright © 2021 by the Society for Music Theory. All rights reserved.

[1] Copyrights for individual items published in *Music Theory Online (MTO)* are held by their authors. Items appearing in *MTO* may be saved and stored in electronic or paper form, and may be shared among individuals for purposes of scholarly research or discussion, but may *not* be republished in any form, electronic or print, without prior, written permission from the author(s), and advance notification of the editors of *MTO*.

[2] Any redistributed form of items published in *MTO* must include the following information in a form appropriate to the medium in which the items are to appear:

This item appeared in *Music Theory Online* in [VOLUME #, ISSUE #] on [DAY/MONTH/YEAR]. It was authored by [FULL NAME, EMAIL ADDRESS], with whose written permission it is reprinted here.

[3] Libraries may archive issues of *MTO* in electronic or paper form for public access so long as each issue is stored in its entirety, and no access fee is charged. Exceptions to these requirements must be approved in writing by the editors of *MTO*, who will act in accordance with the decisions of the Society for Music Theory.

This document and all portions thereof are protected by U.S. and international copyright laws. Material contained herein may be copied and/or distributed for research purposes only.

Prepared by Sam Reenan, Editorial Assistant

The logo for the Society for Music Theory (SMT) features the letters 'SMT' in a large, bold, serif font. The letters are white with a black outline and are set against a background of horizontal lines that create a sense of motion or sound waves. Below the 'SMT' text, the words 'SOCIETY FOR MUSIC THEORY' are written in a smaller, black, sans-serif font.

SOCIETY FOR MUSIC THEORY