A Crosslinguistic Investigation on the Correlation between Functional Load of Tone and Tone-Melody Correspondence

Natsumi Taniguchi (natsumit@usc.edu)
Department of Linguistics, University of Southern California, Los Angeles, CA USA 90089

Keywords: text-setting; tone-melody correspondence; music; Japanese; pitch accent; informativity; functional load

Introduction

Text-setting, or the lining up of lyrics to a melody, is a human behavior that exposes interactions between linguistic and musical knowledge. While following the rules of music, such as being on the pitch and in sync with the rhythm, text-setting reveals that lyricists and singers also use their linguistic knowledge in writing or singing lyrics to a melody. In linguistics, numerous studies have utilized this property to investigate phonological or phonetic questions (e.g. Hayes & Kaun, 1996; McPherson, 2018; McPherson & Ryan, 2018; Starr & Shih, 2017; Wee, 2007).

One way to understand text-setting phenomena is by considering the phonetic perspective. In a song that has lyrics, musical and linguistic pressures are competing over control of the phonetic dimensions of speech/song production. These phonetic dimensions include fundamental frequency (F0), duration, amplitude, vowel quality, aspiration, voicing, and everything else that plays a role in singing. In essence, all of the phonetic components are tools that both the musical pressures and linguistic pressures can utilize to fulfill each of their goals. Since the goals are different between the musical and linguistic pressures, the phonetic components become the target of competition between these two pressures.

Musical pressures may include aesthetic goals, such as emphasis, embellishments, or timbre manipulation, and more basic structural constraints, such as keeping a beat or repeating a phrase to keep to an overall structure. In the case of translated songs, faithfulness to the original melody may be another pressure. Linguistic pressures, on the other hand, are about faithfulness to the lyrics. If there is a conscious (or subconscious) effort towards making sure that the listener can accurately understand the lyrics, which in many cases there probably is, then the composer/performer likely utilizes the phonetic dimensions in such a way that is crucial for accurately understanding words and phrases.

Given this framing of text-setting, the following question can be asked: If a phonological contrast is more useful for accurately recognizing speech in a language (in other words, reducing confusability in speech in a language), then is it more likely to be maintained in text-setting? The current study explores a narrower version of this question focusing on tones: If tone is more important for reducing confusability in speech in language A than in language B, then would tone be more reflected in songs of language A than in songs of language B?

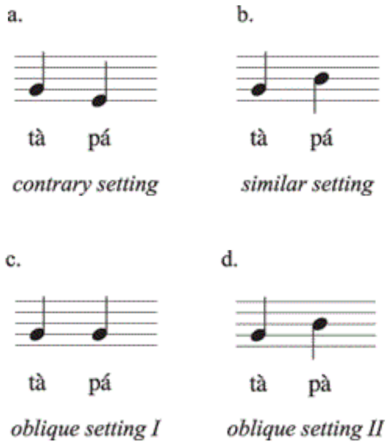

Figure 1: Bigram categories as proposed in Ladd and Kirby (2020)

Tone-melody correspondence

The reflection, or lack thereof, of tone in songs has been investigated often under the name of tone-melody correspondence. An important finding has been that linguistic tones do not correspond to individual musical notes (absolute pitches or pitch classes) in systematic ways. Rather, matching the direction of pitch change in the tones and the melody over multiple syllables/musical notes has often shown positive results. Given this property of tone-melody correspondence, Ladd and Kirby (2020) suggest a descriptive framework for looking at pairs of tones/musical notes, as illustrated in Figure 1. Repurposing terminology from classical Western music theory, contrary setting refers to a consecutive pair of tones/notes, or a bigram, where the tones move in one direction (up or down) while the melody moves in the other direction (down or up). Similar setting refers to a bigram where the tones and notes move in the same direction (up, down, or no movement), regardless of interval. Lastly, oblique setting is a bigram where either the tone or melody moves (up or down) while the other stays level.

Ladd & Kirby (2020) report that the overall pattern found in tone-melody correspondence studies has been 'avoid contrary settings.' In other words, decreasing the number of contrary settings is prioritized over increasing the number of similar settings, and oblique settings are freely allowed in some cases. The pattern is exemplified in numerous tone languages across many song cultures, including traditional Zulu dance songs and Xhosa songs (Rycroft, 1959), traditional and contemporary/popular Thai songs (Ketkaew

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

& Pittayaporn, 2014, 2015; List, 1961), and Tommo So women's folk songs (McPherson & Ryan, 2018), indicating that tone-melody correspondence cannot be attributed to particular cultures or areas. Rather, it seems to be a common characteristic of songs in languages that utilize linguistic tone in one way or another.

Neurophysiological data provides a possible explanation for this pattern. Sadakata et al. (2020) is an EEG study that measures the N2b and N400 components after listening to disyllabic Mandarin words produced on a melody that is either in similar or contrary setting with the tones. In speech, the N2b is often said to be modulated by phonological deviations as well as early stages of semantic processing. The N400 component is generally said to reflect lexical access and semantic integration, as well as phonological processing. The study found that, at least for rising tonal contours, contrary setting triggered a significant N2b response and a marginally significant N400 response, suggesting that processing concurrent incongruent pitch contours (one acoustic, one phonological and abstract) demands higher semantic and phonological processing load. It makes sense that lyricists would avoid increasing processing load, especially in genres that are meant for easy listening, such as popular music, or that are centered around the lyrics.

Cantonese has been frequently shown to have an especially high percentage of similar settings (Chan, 1987b, 1987a; W. Ho Vincie, 2010; Kirby, 2022), as high as 90% (Wong & Diehl, 2002). Interestingly, Mandarin has been the opposite. Kirby (2022) reports that 36% of bigrams were similar settings and 25% were contrary in Mandarin popular songs of the 2000s. The paper subsequently tests and discusses several possible reasons for this discrepancy between Cantonese and Mandarin, two Chinese languages with a complex tonal inventory, and suggests that the proportion of contour tones within the tonal inventory may be a factor.

Schellenberg (2013) also speculates on the difference between Cantonese and Mandarin pop songs in their treatment of linguistic tone. He suggests functional load of tones as a possible linguistic factor, noting the difference in tone inventory size (6-9 in Cantonese vs. 4 in Mandarin), proportion of monosyllabic words in the languages (41.94% in Cantonese according to Qian et al. (2004) and 9.53% in Mandarin according to Chen et al. (1993)) and tonal neighborhood density. The current study is a quantitative investigation of this issue. Functional load, or, intuitively, the work that a phonological contrast does in maintaining meaning contrasts within a language, was used to quantify the usefulness of tone in reducing confusability in speech and test the question of whether this correlates with the degree of tone-melody correspondence in a language.

In addition to the two Chinese languages, a non-canonical tone language was newly analyzed as well to bring a new perspective: Tokyo Japanese (hereon Japanese). Japanese is often called a 'pitch accent' language[1]. It has minimal pairs of words identical in segmental material that are distinguished by their pitch contours over the whole word, meaning pitch is used contrastively. However, the pitch contour is generally rule-governed and predictable excepting the existence and position of the high-low tonal sequence, or the accent. A minimal triplet is shown below in (1) as an example. Each tone ("H" for high, "L" for low) is associated with a mora (mora boundaries are marked with "."). The first two words are distinguished by the position of the high-low tone sequence (marked with an apostrophe) and the third by the absence of a high-low tone sequence. Note that the tone contours are severely restricted in variation, where sequences like HLH, LLH, or LLL are not possible[2].

(1) a. /ha.shi+ga/ 'chopsticks + NOM'
H' L L
b. /ha.shi+ga/ 'bridge + NOM'
L H' L
c. /ha.shi+ga/ 'edge + NOM'
L H H

This word-level specification of tone patterns with limited variation contrasts starkly with more canonical tone languages like Cantonese, where each syllable (usually a morpheme) is associated with one tone out of a multitude of tone identities. Consequently, Japanese is a unique ground for tone-melody correspondence that can expand our knowledge on this phenomenon. Because there is contrastive use of pitch, one might expect there to be some tone-melody correspondence. However, many of the tones are rule-governed within a word, and that may cause lower levels of correspondence than, for example, Cantonese.

The literature has mixed results. Some report that pitch accent does not play a role in Japanese text-setting (Hayes & Swiger, 1995; Kubozono & Mizoguchi, 2023) and focus on syllable weight and whether syllables or moras are matched to musical notes (Starr & Shih, 2017 analyzes this in-depth). Cho (2017) and Taniguchi (2023) report otherwise. Cho (2017) investigates tone/melody bigrams in traditional Japanese children's songs and finds that 54.15% of bigrams were of similar setting, 40.08% were of oblique setting, and 5.77% were of contrary setting. She notes that these numbers are less extreme than those found for Cantonese or Vietnamese, but they differ significantly from chance in the direction of more correspondence. Taniguchi (2023) conducts a perception study where participants are asked to judge the naturalness of various text-setting patterns in the Happy Birthday song's name-inserting portion. It is reported that text-setting patterns that matched the accentual high-low tone sequence with the falling melody (similar setting) were judged more natural than text-setting patterns that did not match tone and melody (contrary and oblique settings). It is

[1] In this paper, the word 'tone' will be used to mean 'contrastive use of pitch' so that it includes Japanese pitch accent. The discussion of whether we should define 'tone' in that way is beyond the scope of this study.

[2] See Kawahara (2015) for an overview of the Japanese pitch accent system.

worth saying that Kubozono & Mizoguchi (2023) report that their production experiment of the Happy Birthday song did not show any influences of pitch accent. Given this, the effect of pitch accent may be subtle and require sensitive enough experiment paradigms. In summary, tone-melody correspondence seems to exist in some datasets, but at lower levels than in more canonical tone languages.

The present study further explores tone-melody correspondence in Tokyo Japanese by looking at a larger and translated dataset of Disney animation movie songs. The songs being translated indicates that there were heavy restrictions in the writing of these Japanese lyrics. The melody cannot be largely altered, limiting the translators to manipulating only the lyrics to increase tone-melody correspondence (if that is one of their conscious or unconscious goals). The overall message of the song is predetermined as well. If tone-melody correspondence shows up in this kind of restricted environment, the role of pitch accent in Japanese text-setting may be more robust than previously thought.

Functional load of tones

Functional load, a concept born in the 20th century from the Prague School (e.g. Mathesius, 1929; Trubetzkoy, 1939), is, intuitively, how much work a phonological contrast is doing for meaning contrasts within a language. Many linguists have explored and supported its possible relation to the diachronic loss of a phonological contrast (e.g. Wedel et al., 2013), overall suggesting its relevance to speakers' phonetic or phonological behaviors. The present study utilizes functional load as a proxy for how much work tones do for accurately understanding speech, as meaning contrasts are crucial for accurately understanding speech.

The functional load of tone in Cantonese and Mandarin have been calculated before using Surendran & Niyogi's (2003) entropy-based methodology twice. Both studies found that Mandarin has similar functional loads between tone and vowels (Oh et al., 2013; Surendran & Levow, 2004). The same was found for Cantonese in Oh et al. (2013). The functional load of tones in Japanese has not been calculated before (to my knowledge), but many have predicted that it would be low based on minimal pair counts (e.g. Matsuo et al., 1961; Sibata & Shibata, 1990). Word recognition experiments also suggest that tones are of lower importance in perception compared to segments. Katsuda & Steffman (2022) conduct a word recognition task with primes, where primes were manipulated to match the target word in just pitch accent, just segments, or both tone and segments (an identity prime). The authors find that pitch accent and segments in Japanese have asymmetric importance in word recognition, where pitch accent priming is only facilitatory when accompanied by segmental priming. Pitch accent-matching primes not accompanied by segmental material of the target are detrimental to word recognition. These results suggest that pitch accent in Japanese has a facilitatory role for word recognition in a minimal (compared to segments) but non-zero way.

Interestingly, Mandarin word recognition experiments show similar results to Japanese. Sereno & Lee (2015) conduct a word recognition experiment with a design similar to Katsuda & Steffman (2022), and they find that tonal information was only facilitatory when accompanied by segmental information and that tonal information unaccompanied with segmental information was inhibitory compared to control conditions. These results can be interpreted to suggest that the functional load of tones in Mandarin may be as low as Japanese.

Methodology

Tone-melody correspondence data and functional load of tone for Mandarin, Cantonese, and Japanese were collected and compared to see if they correlate.

Tone-melody correspondence

For Mandarin and Cantonese, data annotated by Kirby (2022) was used. Setting proportions were recalculated by the author. Each language consisted of 5 pop songs from the 2000s and 2010s, resulting in 1644 and 1720 bigram tokens for Mandarin and Cantonese respectively. For Japanese, 46 songs sung by characters in 12 animation movies released from Walt Disney Pictures were coded for tone-melody bigrams, resulting in 6458 bigram tokens. The movie release dates in Japan range from 1991 to 2014.

For the newly coded Japanese songs, the following procedure was taken. The songs' melody and lyrics were first transcribed by musical note. Melodic transcriptions were based on existing scores, but any alterations made by the Japanese singers were ensured to be transcribed correctly. Bigrams that spanned over musical phrase boundaries were not included. Additionally, all spoken portions were disregarded. Linguistic tones of the lyrics were transcribed based on the author's intuitions as a native Tokyo Japanese speaker, except for any words that came up in dialogue within the clip of the song, in which case the character's pronunciation was prioritized. To keep the tone bigrams free of intonational intuitions, judgments were made at the lexical level (including compounds and verbal constructions) and all sentence final particles were excluded. Once all musical notes and linguistic tones were transcribed, the direction of pitch change was categorized as rising, falling, or level in each melodic or tonal bigram. Corresponding bigrams of melody and tone were then categorized into the 9 possible combinations.

Functional load

The methodology for calculating functional load has been debated over the years, ranging from simple minimal pair counts to system entropy-based methods. Minimal pair counts are simple and have been argued to predict diachronic phoneme contrast loss better than other measures (Wedel et al., 2013). System entropy-based methods, on the other hand, can take token frequency into account, which minimal pair counts in most methods do not. In the present study, a system entropy-based methodology was chosen for this reason,

although a comparison with minimal pair count results may reveal whether token or type counts are more relevant to how lyricists contextualize tone in music.

In particular, the methodology developed in Surendran & Niyogi (2003) was utilized. In essence, this method calculates functional load as how much work tones do in the (bigram) predictability of morphemes in a language's speech. The conversational speech corpora, Hong Kong Cantonese Corpus (Luke & Wong, 2015), Taiwan-Beijing Mandarin Youtube Corpus (Huang, n.d.), and Japanese Topic-Oriented Conversation Corpus (Nakamata et al., 2021) were used. Each had a morpheme count of 191863, 1012448, and 47394.

Predictions

Given the existing literature, the following predictions can be made. First, for tone-melody correspondence, Cantonese will have a high proportion of similar settings and a low proportion of contrary settings compared to the other two languages. Kirby's (2022) results show that Cantonese songs do have these characteristics compared to Mandarin at least. Mandarin and Japanese will have fairly comparable numbers.

Second, for the functional load of tone, Cantonese and Mandarin will have a fairly high functional load of tone within the language. Mandarin may have a slightly lower number. Japanese will show a fairly low functional load of tone based on the minimal pair count studies. If the word recognition experiments are taken into account, however, one might expect Mandarin and Japanese to have a similar functional load for tone compared to that of segments.

Overall, the prediction is that the degree of tone-melody correspondence and the functional load of tones correlate between the three languages.

Results

Tone-melody correspondence

The stacked bar graph below in Figure 2 summarizes the proportions of similar, contrary, oblique (level melody), and oblique (level tone) settings in all three languages. Each language has two bars, the left reporting the observed proportions and the right reporting the expected proportions, calculated by multiplying the proportions of rising/falling/level melody with the proportions of rising/falling/level tonal bigrams within each corpus.

As expected, Cantonese has a high percentage of similar settings (77.33%) and a low percentage of contrary settings (4.88%). Oblique (level melody) settings also have a low percentage at 3.26%, suggesting that Cantonese songs increase similar settings even when the melody is level. The observed proportions are clearly different from the expected proportions towards the direction of more tone-melody correspondence. A chi-squared goodness of fit test showed statistical significance $\chi^2(3, N=1720) = 1520.883$, $p = < 0.001$. Both Mandarin and Japanese have relatively low tone-melody correspondence compared to Cantonese, as predicted. Chi-squared goodness of fit tests for both languages between the expected and observed proportions are significant (Mandarin: $\chi^2(3, N=1644) = 14.455$, $p = 0.002347 < 0.01$, Japanese: $\chi^2(3, N=8365) = 46.936$, $p < 0.001$). For both languages, there are fewer contrary settings and more similar settings in the observed than the expected. Therefore, it can be said that there is a statistically significant degree of tone-melody correspondence in both Mandarin and Japanese, at least in these corpora.

In summary, Cantonese had the highest degree of correspondence. Mandarin and Japanese seem to be fairly comparable in their numbers, although Japanese may have a slight edge.

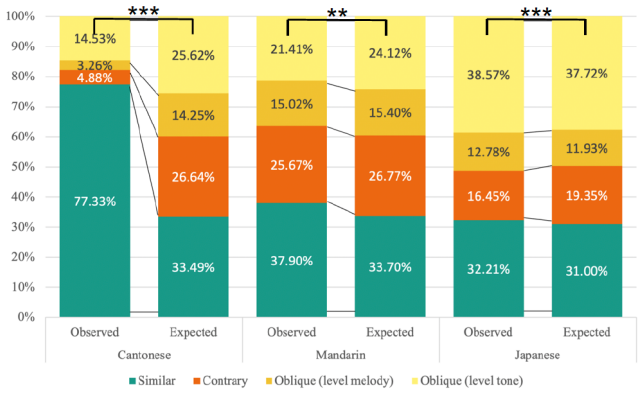

Figure 2: Tone-melody correspondence proportions

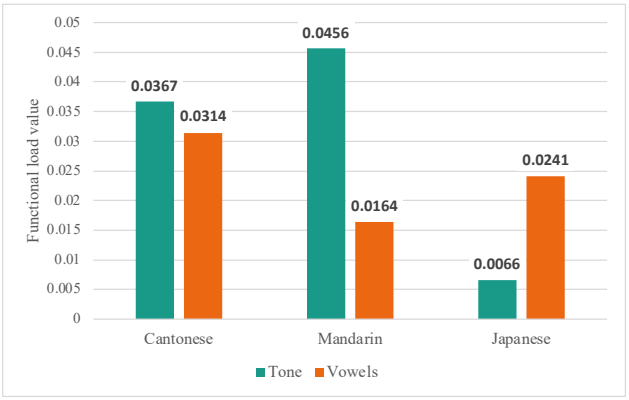

Figure 3: Functional load of tone and vowels

Functional load

Figure 3 summarizes the functional load values for tone and vowels in the three languages. For Cantonese, the functional load of tone (0.0367) and vowels (0.0314) are comparable, replicating Oh et al. (2013). For Mandarin, the functional load for tone (0.0456) is much higher than that for vowels (0.0164). This is a little unexpected, given that both Surendran & Levow (2004) and Oh et al. (2013) reported comparable numbers for tone and vowels. For Japanese, the functional load for tone (0.0066) was fairly low compared to the functional load for vowels (0.0241), as expected from the minimal pair count and word recognition literature. Overall, the results show that Mandarin tones have a higher functional

load within the language than Cantonese tones, and Japanese tones have the lowest functional load.

Overall comparison

Tone-melody correspondence was highest for Cantonese and comparable between Mandarin and Japanese, although Japanese had a slight edge. The functional load of tone was highest for Mandarin, then Cantonese, and lastly Japanese. These results are visualized together in the scatter plot in Figure 4. The y-axis represents the ratio of similar to contrary settings in each language's dataset. This measure takes into account both whether the language maximized similar settings or avoided contrary settings. If our hypothesis that the functional load of tone would correlate with the degree of tone-melody correspondence were supported, the data points in the scatterplot would roughly line up from the bottom left corner to the top right corner. However, it can be seen that Mandarin is in the bottom right, indicating that only Cantonese and Japanese behave as hypothesized.

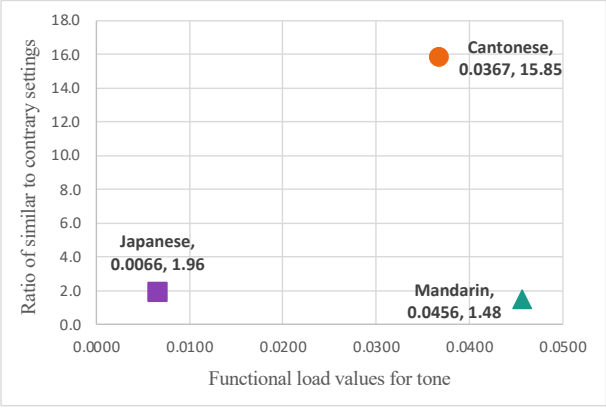

Figure 4: Comparison of functional load values for tone and degree of tone-melody correspondence in Cantonese, Mandarin, and Japanese

Discussion

The functional load values of tone and tone-melody correspondence degrees calculated in this study did not correlate in the three languages. However, the Cantonese and Japanese data points in Figure 4 are positioned as hypothesized. Mandarin seems to be the odd one out. This section will discuss the possible reasons why Mandarin songs only show a small degree of tone-melody correspondence even when it has a high functional load for tone.

Phonological differences between tonal systems

Kirby (2022) tests three phonological differences between Cantonese and Mandarin (proportion of disyllabic inventory, neutral tones, and tone sandhi differences) and finds that none have a significant effect on the proportions of similar, contrary, and oblique settings. As an alternative possibility, he proposes that the proportion of contour tones within tonal inventories plays a role in tone-melody correspondence. The idea is that contour tones, especially those that are produced with a relatively large F0 movement, are not ideal for singing, where in most cases, one syllable must be produced as one musical note. Mandarin has 2 (or 3 at utterance final position) contour tones within a 4-tone system, while Cantonese has at most 3 contour tones within a 6-tone system. Kirby postulates that Mandarin has a harder time achieving tone-melody correspondence because of the higher proportion of contour tones. However, this hypothesis has counterexamples such as Zulu and Xhosa traditional songs that show high levels of tone-melody correspondence (Rycroft, 1959, 1979).

Phonetic realization of tones in Mandarin

The phonetic correlates for Mandarin tones are not just F0. It has been reported that in both production and perception, duration and intensity also consistently play a role in Mandarin tones. Whalen & Xu (1992) conduct a series of perception experiments and find that, although it is clear that F0 is the main source of information for differentiating tones, amplitude contours also have significant information for the listeners. Ho (1976) among others finds that duration also differs between tones in production. Specifically, in isolated monosyllable productions, T3 (low falling-rising) has the longest duration and T4 (high-falling) the shortest, with T1 and T2 falling in between. Although duration and amplitude seem to be secondary cues in regular speech, they can play a crucial role in tonal contrast perception in restricted environments like whispered speech (Gao, 2002). Therefore, it would not be surprising to find that tones with longer duration (T3) tend to be allotted longer durations in songs as well, since singing is an environment that restricts F0 use due to musical pressures, and this could be allowing for less tone-melody correspondence (i.e. F0-matching). Zhang et al. (2024) tests this hypothesis with a production experiment and found that duration cues of regular speech were partially replicated in singing. They report that T4 was the shortest in singing as well, but T1 duration was longer than T3 in singing. The study concludes that speakers may partially rely on the secondary cue of duration in singing. Hence, it is not conclusive whether duration is reflected in songs consistently and whether that is causing lower degrees of tone-melody correspondence. As for Cantonese, there are also studies that show duration differences between tones in production. Unless the durational cues in Cantonese are much weaker or have much less information than in Mandarin, this phonetic correlate explanation may not be able to explain the discrepancy between the two languages' tone-melody correspondence degrees.

Another possible reason is related to how Mandarin tones are realized in natural speech inside phrases and sentences. In its conclusion, Chan (1987b) cites Tseng (1981) which reports that F0 productions of tones in natural speech Mandarin are only 40.17% of the time faithful to the phonological prediction of what phonetic values they would be because of intonational contours. Chan draws a parallel between this domination of intonational contours over tones with the domination of melody over tones. This observation

suggests a possibility that is crucial to our research question; Perhaps in Mandarin, consistent production of F0 contours for tones is not that critical for accurately understanding natural speech.

Effect of musical genre and culture

Lastly, there is always the possibility that transmitting lyrics is not a priority in the musical culture of Mandarin. Perhaps Cantonese pop music happened to care about matching tones and melodies while Mandarin pop music did not. Differences between musical genres within one language culture have been documented before in studies such as (List, 1961). A difference this large between Mandarin and Cantonese pop cultures is peculiar, however, given the shared history between the two musical cultures (Schellenberg, 2013)[3].

Conclusion and Future Directions

In this study, the question of whether a phonological feature's usefulness for reducing confusability in speech in a language correlates with it being reflected in songs was investigated, focusing on tones as the phonological feature and tone-melody correspondence as the way for tones to be reflected in songs. The functional load of tone and degree of tone-melody correspondence in Cantonese, Mandarin, and Japanese were calculated and compared. The calculated values did not correlate, and, therefore, our hypothesis was not supported. However, Cantonese and Japanese did behave as hypothesized. It may be the case that Mandarin has some unique property that makes it deviate from the hypothesized pattern. Unrelated to the main research question, the data analyzed in this study shows that tone-melody correspondence in Japanese exists in restricted domains such as translated songs although at a lower degree than in less restricted domains such as traditional children's songs. In a broader sense, the information theoretic approach to text-setting explored in this study provides a novel window into investigating how language and music, two different cognitive modules, interact and produce certain behavioral trends.

There are several limitations to this study. First, the tone-melody correspondence data differed between languages. Keeping this consistent would be ideal to control for genre and restrictedness of lyric writing.

Second, we still do not know what is causing Mandarin to have a low degree of tone-melody correspondence despite its high functional load of tone. As discussed, several different approaches can be made to further investigate this issue that focus on linguistic characteristics of Mandarin. Another direction may be to simply add more datapoints to Figure 4 by expanding the current study to more languages. It would be especially interesting to analyze Bantu languages such as Xhosa and Zulu which have much smaller tonal systems compared to the Chinese varieties, but still seem to show high degrees of tone-melody correspondence. If there is an overall correlation between the functional load of tone and the degree of tone-melody correspondence with the addition of these languages, this would suggest that the significance of pitch in language affects how musicians produce (and likely perceive) lyrics in a musical context.

Third, this study is very coarse-grained in that it investigates the language-level correlation between functional load of tone and its reflection in music. A more insightful way to deepen our understanding of how linguistic and musical cognition work together in text-setting may be to analyze it at the bigram level. This may investigate how linguistic factors such as functional load between individual tonal contrasts, word boundaries, pitch accent nucleus (if Japanese), and local predictability of the segments, and musical factors such as metrical strength, magnitude of melodic transition, and phrase boundaries affect the tone-melody correspondence of a particular bigram.

[3] The diachronic collapse of tonal contrasts in Mandarin may have to do with the current musical culture in relation to tone-melody correspondence (p.c. Bruce Hayes).

Acknowledgments

Thank you to my advisor Stephanie Shih in guiding me through this project. I am also very grateful for the comments from Louis Goldstein, Canaan Breiss, and all others who were at presentations throughout the process.

References

Chan, M. (1987a). Tone and Melody in Cantonese. *Annual Meeting of the Berkeley Linguistics Society*, 26–37. https://doi.org/10.3765/bls.v13i0.1828

Chan, M. (1987b). Tone and melody interaction in Cantonese and Mandarin songs. *UCLA Working Papers in Phonetics*, *68*, 132–169.

Chen, C.-Y., Tseng, S.-F., Huang, C.-R., & Chen, K.-J. (1993). *Some distributional properties of Mandarin Chinese: A study based on the Academia Sinica Corpus*.

Cho, S. (2017). Text alignment in Japanese children's song. *University of Pennsylvania Working Papers in Linguistics*, *23*(1), 5.

Gao, M. (2002). *Tones in whispered Chinese: Articulatory features and perceptual cues*.

Hayes, B., & Kaun, A. (1996). The role of phonological phrasing in sung and chanted verse. *The Linguistic Review*, *13*(3–4). https://doi.org/10.1515/tlir.1996.13.3-4.243

Hayes, B., & Swiger, T. (1995). *TWO JAPANESE CHILDREN'S SONGS*.

Ho, A. T. (1976). The Acoustic Variation of Mandarin Tones. *Phonetica*, *33*(5). https://doi.org/10.1159/000259792

Ho, W., Vincie. (2010). *A phonological study of the tone-melody correspondence in Cantonese pop music* (pp. 991035825399703414, b50899752) [Doctor of Philosophy, The University of Hong Kong]. https://doi.org/10.5353/th_b5089975

Huang, P.-H. (n.d.). *Mandarin classifier choice and information-theoretic constraints. [Unpublished manuscript]*.

Katsuda, H., & Steffman, J. (2022). Asymmetrical roles of segment and pitch accent in Japanese spoken word recognition. *JASA Express Letters*, *2*(6), 065201. https://doi.org/10.1121/10.0011573

Kawahara, S. (2015). Chapter 12 The phonology of Japanese accent. *Handbook of Japanese Phonetics and Phonology*.

Ketkaew, C., & Pittayaporn, P. (2014). *Mapping between lexical tones and musical notes in Thai pop songs*. 160–169.

Ketkaew, C., & Pittayaporn, P. (2015). *DO NOTE VALUES AFFECT PARALLELISM BETWEEN LEXICAL TONES AND MUSICAL NOTES IN THAI POP SONGS?*

Kirby, J. P. (2022). Comparative tonal text-setting in Mandarin and Cantonese popular song. *Proceedings of CLS*, *58*.

Kubozono, H., & Mizoguchi, A. (2023). *TEXT-TO-TUNE ALIGNMENT IN JAPANESE "HAPPY BIRTHDAY TO YOU."*

Ladd, D. R., & Kirby, J. (2020). Tone–Melody Matching in tone-Language Singing. In C. Gussenhhoven & A. Chen (Eds.), *The Oxford Handbook of Language Prosody* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198832232.013.47

List, G. (1961). Speech Melody and Song Melody in Central Thailand. *Ethnomusicology*, *5*(1), 16–32. https://doi.org/10.2307/924305

Luke, K. K., & Wong, M. L. (2015). The Hong Kong Cantonese corpus: Design and uses. *Journal of Chinese Linguistics Monograph Series*, *25*.

Mathesius, V. (1929). La structure phonologique du lexique du tchèque moderne. *Travaux Du Cercle Linguistique de Prague*, *1*, 67–84.

Matsuo O., Itikawa T., Okubo A., & Tanaka A. (1961). *douon-go no kenkyuu "Japanese homonymy and its problems."* NINJAL. https://doi.org/10.15084/00001232

McPherson, L. (2018). The role of music in documenting phonological grammar: Two case studies from West Africa. *Proceedings of the Annual Meetings on Phonology*. https://doi.org/10.3765/amp.v7i0.4493

McPherson, L., & Ryan, K. M. (2018). Tone-tune association in Tommo So (Dogon) folk songs. *Language*, 119–156.

Nakamata, N., Ota, Y., Kato, E., Sawada, H., Shimizu, Y., & Mori, A. (2021). J-TOCC: Japanese Topic-Oriented Conversation Corpus. *Mathematical Linguistics*, *33*(1). https://doi.org/10.24701/mathling.33.1_11

Oh, Y. M., Pellegrino, F., Coupé, C., & Marsico, E. (2013). Cross-language comparison of functional load for vowels, consonants, and tones. *Interspeech 2013*, 3032–3036. https://doi.org/10.21437/Interspeech.2013-662

Qian, Y., Lee, T., & Soong, F. K. (2004). Use of tone information in continuous Cantonese speech recognition. *Speech Prosody 2004*. https://doi.org/10.21437/SpeechProsody.2004-135

Rycroft, D. (1959). African Music in Johannesburg: African and Non-African Features. *Journal of the International Folk Music Council*, *11*, 25–30. https://doi.org/10.2307/834851

Rycroft, D. (1979). The relationships between speech-tone and melody in Southern African music. *South African Music Encyclopedia*, *2*.

Sadakata, M., Weidema, J. L., & Honing, H. (2020). Parallel pitch processing in speech and melody: A study of the interference of musical melody on lexical pitch perception in speakers of Mandarin. *PLOS ONE*, *15*(3), e0229109. https://doi.org/10.1371/journal.pone.0229109

Sereno, J. A., & Lee, H. (2015). The Contribution of Segmental and Tonal Information in Mandarin Spoken Word Processing. *Language and Speech*, *58*(2), 131–151. https://doi.org/10.1177/0023830914522956

Sibata, T., & Shibata, R. (1990). "akusento wa douongo wo dono teido benbetsu shiuruka [To what extent can accent distinguish homophones?]." *Mathematical Linguistics: 17*(7), p317-327.

Starr, R. L., & Shih, S. S. (2017). The syllable as a prosodic unit in Japanese lexical strata: Evidence from text-setting. *Glossa: A Journal of General Linguistics*, *2*(1). https://doi.org/10.5334/gjgl.355

Surendran, D., & Levow, G.-A. (2004). The functional load of tone in Mandarin is as high as that of vowels. *Speech Prosody 2004*, 99–102. https://doi.org/10.21437/SpeechProsody.2004-23

Surendran, D., & Niyogi, P. (2003). *Measuring the Functional Load of Phonological Contrasts* (arXiv:cs/0311036). arXiv. https://doi.org/10.48550/arXiv.cs/0311036

Taniguchi, N. (2023). Exploring text-to-tune alignment of Japanese special moras in "happy birthday to you." *ICU Working Papers in Linguistics (ICUWPL)*, *24*, 81–92.

Trubetzkoy, N. S. (1939). Grundzüge der Phonologie (Travaux du cercle linguistique de Prague 7). *Prague: Cercle Linguistique de Prague*.

Tseng, C.-Y. (1981). *An Acoustic Phonetic Study on Tones in Mandarin Chinese* [Ph.D.]. Brown University.

Wedel, A., Kaplan, A., & Jackson, S. (2013). High functional load inhibits phonological contrast loss: A corpus study. *Cognition*, *128*(2), 179–186. https://doi.org/10.1016/j.cognition.2013.03.002

Wee, L. H. (2007). Unraveling the relation between Mandarin tones and musical melody. *Journal of Chinese Linguistics*, *35*(1), 128.

Weidema, J. L., Roncaglia-Denissen, M. P., & Honing, H. (2016). Top–Down Modulation on the Perception and Categorization of Identical Pitch Contours in Speech and Music. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00817

Whalen, D. H., & Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica*, *49*(1), 25–47.

Wong, P. C., & Diehl, R. L. (2002). How can the lyrics of a song in a tone language be understood? *Psychology of Music*, *30*(2), 202–209.

Zhang, Q., Zhu, L., & Jiang, X. (2024). Can we sing the tones of a tonal language? The duration of Mandarin tones under music context. *Speech Prosody 2024*, 344–348. https://doi.org/10.21437/SpeechProsody.2024-70