An analysis of rhythm in Japanese and English popular music

Makiko Sadakata*,

Peter Desain*, Henkjan Honing**

Aniruddh D. Patel***, John R. Iversen***

*NICI, University of Nijmegen
**ILLC/ Music Department, University of Amsterdam
***The Neurosciences Institute, San Diego, California
m.sadakata@nici.kun.nl

Abstract: Recently, there has been evidence that the rhythm in English and French non-vocal musical themes are significantly different in their contrastiveness of successive durations in the same manner as those of spoken language, suggesting that a composer's native language exerts an influence on the music composed (Patel & Daniele, 2003). This has been shown using normalized Pairwise Variability Index (nPVI; Grabe & Low, 2002). Would such an effect of a composer's native language also exist for music written with lyrics? Or would the language of the lyrics exert a stronger rhythmic influence on the music? In the present study, we compare rhythm in popular music with Japanese and English lyrics. The preliminary result, for music written by Japanese composers, showed a small but significant difference between them, suggesting successive durations in Japanese popular music are less contrastive than in English. The effect is in accordance with the differences between spoken Japanese and English, and suggests that the language of the lyrics with which the music is composed exerts an influence on the rhythmic structure of the music, despite the composer's native language.

Keywords: speech rhythm, musical rhythm

1. Introduction

Cultural impact on music has been a focus of attention in musicology for a long time. What aspects of composers', performers' and listeners' culture affect how music is written, performed or perceived? More specifically, many efforts have been made to associate different aspects of musical behavior with the characteristics of language. Lerdahl & Jackendoff (1983)'s influential work, *A Generative Theory of Tonal Music*, points at a similarity of the syntactic structure of music and language. In addition to formal studies of structural similarity, there have been numerous empirical investigations of connections between music and language. Since music and language are both conveyed as sequences of sounds organized in time, the study of temporal, or rhythmic, relationships is of potential interest. Several studies have found evidence for relations between rhythmic aspects of music and language. Wenk (1987) showed that certain musical works written by French composer tend to reflect prosodic characteristics of the French language. More recently, Patel & Daniele (2003) applied measures developed to describe rhythmic differences between languages to classical musical themes, and found that music composed by French and English composers of a certain era reflected the rhythmic differences between the

composer's native languages. The themes compared were carefully selected to exclude vocal themes with lyric content to exclude the possible influence of the language of the lyrics. The current paper extends these past studies to popular music with lyrics. In this situation, both the language of the composer and the language of the lyrics can potentially affect the musical rhythm. We present preliminary results of whether linguistic rhythmic differences are expressed in popular music written with either Japanese or English lyrics.

The measure of rhythm used by Patel and Daniel (2003) is called the "normalized Pairwise Variability Index" (nPVI) and was developed by Grabe & Low (2002) for rhythmic analysis of vocalic durations in speech utterances. The nPVI is a measure of the contrastiveness of successive durations and was applied by Patel and Daniel (2003) to the rhythmic analysis of musical scores. Specifically, they found evidence that composed rhythm in English and French musical themes is significantly different in their contrastiveness of successive durations in the same manner as those of spoken language. The nPVI was devised as a new empirical tool of analyzing speech rhythm in the field of phonetics. The premise behind it is that language can be classified based on the proportion of durations of vocalic and intervocalic intervals and contrastiveness of successive these durations (Ramus et al, 1999; Grabe & Low, 2002). This approach has successfully shown an empirical difference between so-called stress-timed languages and syllable-timed languages, with less contrastiveness of successive vocalic durations for syllable-timed languages. Similarly, several studies (Ramus et al., 1999; Ramus, 2002; Warner & Arai, 2001) revealed distinct rhythmic features of Japanese, which belongs to third group of so-called mora-timed languages. Table 1 summarizes the basic characteristics of each language class regarding relative values of vocalic nPVI and intervocalic rPVI (raw Pairwise Variability Index) represented as Figure 2 in the study by Ramus (2002). Japanese was similar to syllable-timed languages in showing less contrastiveness of successive vocalic durations, but was distinguished by showing less contrastiveness of successive intervocalic durations than the other two language classes.

In this study we expect to find more evidence of linguistic impact on musical composition using Japanese (mora-timed) and English (stress-timed) music as a case study of the method by Patel & Daniele (2003). The method relates contrastiveness of vocalic durations in speech to that of duration of intervals in music, which means that we cannot provide Japanese specific feature, as it doesn't take intervocalic intervals into account. However, as Japanese and French (syllable-timed) have virtually identical nPVI values for speech (Ramus, 2002), in contrast with English, less contrastiveness of successive durations in Japanese music than English music would be expected.

2. Method

The raw Pairwise Variability Index (rPVI) is defined as:

$$rPVI = \frac{100}{m-1} \sum_{k=1}^{m-1} \left| d_k - d_{k+1} \right| \quad (1)$$

where m is the number of intervals and d_k is the duration of the kth interval (Ramus, 1999; Ramus et al., 2002).

Table 1 Summary of characteristics of each language class regarding to contrastiveness of successive intervocalic / vocalic durations, presented as Fig. 2 in Ramus (2002). See formula (1) and (2) for the definitions of rPVI and nPVI.

Language class	Languages	Intervocalic rPVI	Vocalic nPVI
Stress-timed	English, German	High	High
Syllable-timed	Catalan, French, Italian, Polish, Spanish, Catalan	High	Low
Mora-timed	Japanese	Low	Low

Application of nPVI into the analysis of musical (speech) rhythm is defined as:

$$nPVI = \frac{100}{m-1} \sum_{k=1}^{m-1} \left| \frac{d_k - d_{k+1}}{\frac{d_k + d_{k+1}}{2}} \right| \quad (2)$$

where m is the number of intervals and d_k is the duration of the kth interval (Grabe & Low, 2002; Patel & Daniele, 2003). The average duration of two successive intervals normalizes the difference of these two durations.

3. Material

The nPVI was employed to analyze midi files of popular music from the RWC Music Database (Popular Music, Goto et al, 2002). The database contains 80 Japanese and 20 English popular music with lyrics. Note that the English music in this database was written by Japanese composers.

Careful processing of musical data has been done in order to provide a proper sample for nPVI calculation. Durations longer than one bar are omitted from the target, and rhythmic phrase including less than 12 successive notes are removed from the target. POCO (Honing, 1990) was used for the segmentation of musical materials. Sample size was 157 musical phrases (7,129 subsequent durations) for English and 494 (27,341 subsequent durations) phrases for Japanese music, respectively.

4. Result

nPVI values were computed from music phrases and compared between English and Japanese popular music. There was a significant difference between the language groups regarding to nPVI value (Mann-Whitney U test, p=0.0397) in the same direction as the difference between Brigish English & Japanese speech. Fig.1 shows that calculated nPVI value for each language group (English: mean nPVI = 53.0, standard deviation = 14.7, Japanese: mean nPVI = 49.9, standard deviation = 15.4).

Although we expected a large difference in their nPVI because the difference between British English and Japanese in speech shown by Ramus (2002) was quite large (approximately 20 nPVI points), the size of the difference in this English vs. Japanese popular music was rather small (about 3 nPVI points). This 3 nPVI points is also smaller than that of for English vs. French music which Patel & Daniele (2003) obtained (about 6 nPVI points).

5. Discussion

Following an observation by Ohgushi (2002), in Sadakata et al. (In press), it was investigated in how far Japanese and Western musicians differ in the performance of simple patterns. However, in these studies a possible cause of these differences was not discussed. As Japanese musicians are presently exposed to the same amount of western music as western musicians are, it is not likely that the difference in musical

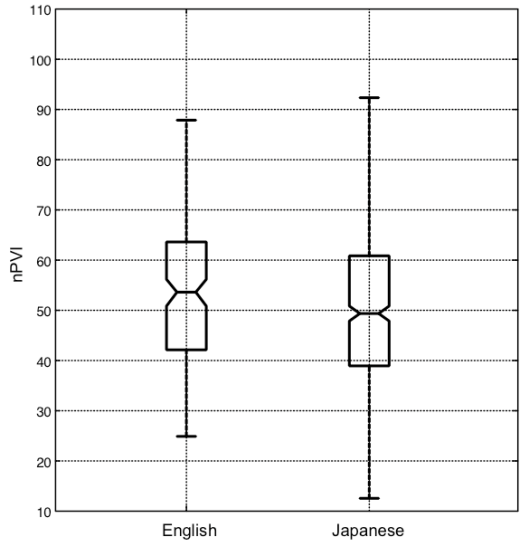

Fig. 1 Musical nPVI values for popular music in English and Japanese. The ends of the box are the 25th and 75th quantiles, the line across the middle of the box identifies the median (English=53.4, Japanese=49.3). Lines that extend from the ends of the box indicate the outermost data point that falls within the distances computed.

cultural tradition is a cause of any difference. However the exposure to the mother tongue is of course a major distinction between western and Japanese musicians and composers. The aim of this study is to broaden the application of nPVI to compare speech and music with the purpose of investigating an influence of the prosody of language on musical rhythm. Preliminary result showed less contrastiveness of successive durations in Japanese than English popular music, which is in accordance with the tendency of spoken language. However difference of nPVI points between groups was smaller than expected at this stage.

This small nPVI difference might be simply because Japanese composer composed both the English and Japanese music. However, Patel & Daniele (2003) showed that the nPVI value for one French composer (syllable-timed) whose style of composition aspired for German music tradition (stress-timed) was extremely high. In this manner, Japanese composer's elaboration of English popular music could have reflected as large nPVI value for English popular music, as they tried to replicate the style of English popular music. The existence of lyrics was also expected to promote differentiation of nPVI points for music from different language groups.

On the other hand, there may be indeed only slight influence of the prosody of a composer's native language on popular music because the rhythmic structure of popular music is (intuitively) simpler than that of classic music, even it is with lyrics. One way of understanding rhythmic structure is to investigate its component (simple rhythmic patterns). Sadakata et al. (submitted) showed that exposure to simple rhythmic patterns to everyday life may be universal regardless of the type of music, by counting up the frequency of occurrence of the two-interval rhythmic patterns (such as 1:2, 1:3 etc.) from quite different musical corpora (Folksongs, National Anthems, Western Classical music). Since the investigation of rhythmic structure of popular music as compared to classical music is an important issue, the same approach could be applied to other musical material. Given these perspectives, further investigation will have to be carried out using English popular music written by English speaking composers.

References

Goto, M., Hashiguchi, H., Nishimura, T., & Oka, R. (2002). RWC Music Database: Popular, Classical, and Jazz Music Database, *Proceedings of the 3rd International Conference on Music Information Retrieval* (pp.287-288), Paris, October.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology 7* (pp. 515–546). Berlin: Mouton de Gruyter.

Honing, H. (1990). POCO: an environment for analyzing, modifying, and generating expression in music. *Proceedings of the International Computer Music Conference*. (pp.364-368). San Francisco: Computer Music Association.

Lerdahl, R., & Jackendoff, F. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Ohgushi, K. (2002). Comparison of Dotted Rhythm Expression between Japanese and Western Pianists. *Proceedings of the 7th International Conference on Music Perception & Cognition*, Sydney, July.

Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, 87:B35-B45.

Ramus, F. (2002). Acoustic correlates of linguistic rhythm: Perspectives. *Proceedings of Speech Prosody* (pp.1 15-120), Aix-en- provence.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 72, 265-292.

Sadakata, M., Desain, P., & Honing, H. (submitted). The relation between rhythm perception and production: towards a Bayesian model.

Sadakata, M., Ohgushi, K., & Desain, P. (in press). A cross-cultural comparison study of the production of simple rhythmic patterns. *Psychology of Music.*

Warner, N., & Arai, T. (2001). The role of the mora in the timing of spontaneous Japanese speech. *Journal of the Acoustical Society of America*, 109, 1 144-1 156.

Wenk, B. J. (1987). Just in time: on speech rhythms in music. *Linguistics*, 25, 969-981.