

TRENDS IN/OVER TIME: RHYTHM IN SPEECH AND MUSICAL MELODY IN 19TH-CENTURY ART SONG

Leigh VanHandel

School of Music
Michigan State University
102 Music Building
East Lansing, MI 48823

ABSTRACT

This paper presents the results of a quantitative study of the relationship between rhythmic characteristics of spoken German and French and the rhythm of musical melody in 19th-century art song. The study used a modified version of the *Normalized Pairwise Variability Index*, or *nPVI*, to measure the amount of variability between successive rhythmic events in the melodies of over 600 songs by 19 French and German composers. The study returned an unexpected result; songs written to texts in the two languages exhibited sharply diverging trends as a function of time through the 19th century. This trend is reflected both in the overall trends and in the trends of individual composers.

1. INTRODUCTION

Rhythm in spoken language is an effect involving the *isochrony*, or the regularity of occurrence, of some type of speech item; the item can be a speech unit (such as a syllable) or a linguistic event (such as accent or stress). Languages may be categorized as *syllable-timed* (for example, French and Spanish) or *stress-timed* (for example, English and German).

In a *stress-timed* language such as German, stressed syllables are perceived to occur regularly, resulting in an expansion or compression of individual syllables in order to accommodate the recurrence of the stresses. By contrast, in a *syllable-timed* language such as French, it is the individual syllables that are perceived to be of nearly equal duration and therefore occurring at regular time intervals.

The linguistic classifications of *syllable-timed* and *stress-timed* are controversial ones. Repeated attempts to empirically confirm isochrony between stresses or syllables have failed [1,5]. However, many linguists agree that the principle of isochrony can underlie the rhythm of a language even if it is not manifested in an experimentally demonstrable fashion [2]. The durational differences between stresses or syllables tend to be well below the threshold of perception, so even though the principle cannot be proven quantitatively, we still perceive the unit to be recurring isochronously. As a result, many linguists now regard isochrony as a primarily perceptual phenomenon, describing it as a “subjective” rather than “objective” isochrony.

Some of the strongest evidence for the perceptual-based theory of linguistic isochrony comes from studies done on language acquisition and categorization in infants which demonstrate that infants are able to discriminate on the basis of rhythm between sentences in their native language and sentences from a language belonging to another rhythmic class [7]. The perceptual processes involved in discriminating language on the basis of rhythm have also been found in tamarin monkeys and most recently, in rats [8].

Recent studies, most notably Grabe & Low [3] and Ramus *et al.* [7], have demonstrated that there are quantitative rhythmic differences between languages historically considered to be stress- and syllable-timed languages. These studies have focused on measuring vowel duration in the two types of languages; they hypothesized that since vowels are primarily responsible for the length of a syllable, an increased variability in vowel length would result in greater variability in syllable length, whereas a language with little variation in vowel length would have little overall variation in syllable length. The model Grabe & Low [3] developed to measure the amount of durational variability in a language is called the *normalized Pairwise Variability Index*, or *nPVI*, defined as:

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

where m is the number of items and d is the duration of the k th item. Grabe and Low’s research demonstrated that because stress-timed languages have greater durational variability between successive vocalic intervals due to compression and expansion of syllables, stress-timed languages will have a higher *nPVI* measurement than a syllable-timed language with lower variability.

2. APPLICATION TO MUSIC

2.1. Previous studies

While the *nPVI* was originally intended to be a measure of the average variation of successive vocalic intervals in spoken language, several studies have used the measure to study the relative durational contrast between successive rhythmic events in music. The following

sections summarize the methodology and results from previous studies.

2.1.1 Patel and Daniele (2003) [6]

Patel and Daniele [6] used the nPVI measure to study fragments from British and French instrumental music. The musical material for Patel and Daniele's study was drawn from the Barlow and Morgenstern *A Dictionary of Musical Themes, 2nd Edition* (1983), and focused on 19th–20th century composers¹ who were native speakers of British English and continental French. The themes were selected using very strict criteria, narrowing the number of themes eligible for the study to a total of 319 themes from six English and ten French composers.

Patel and Daniele's results indicated that the rhythm of instrumental music by British composers demonstrated a higher durational contrast than that of French composers, as expected based on the higher durational contrast in spoken British English than in spoken French.

2.1.2 Huron and Ollen (2003) [4]

Huron and Ollen [4] replicated Patel and Daniele's study, dramatically expanding the sample size. Their results for English and French were consistent with those of Patel and Daniele, although the difference between the mean values was smaller than that found by Patel and Daniele; nonetheless, the results were significant, indicating the hypothesis holds up even with less stringent sampling criteria.

In addition to confirming Patel and Daniele's original hypothesis about British English and French, Huron and Ollen calculated average nPVI values for instrumental music written by composers of other nationalities (results reproduced in Table 1). They noted that results were comparable for related languages such as American English and British English, for Austrian and German, and for syllable-timed Romance languages such as French, Spanish, and Italian.

Nationality	(nPVI)	Standard Error	No. of Themes
American	46.3	1.12	415
Austrian	42.0	0.65	1194
Czech	47.1	1.56	232
English	45.6	0.90	737
French	43.7	0.73	1188
German	42.0	0.59	2006
Hungarian	45.4	1.62	244
Italian	42.7	1.06	529
Polish	45.1	1.74	254
Russian	39.8	0.82	736
Scandinavian*	45.9	1.79	141
Spanish	42.5	1.85	108

* Danish, Swedish, Norwegian (Finnish excluded)

Table 1. Mean nPVI results from Huron and Ollen [4]

¹ Their criterion was that the composer was born in the 19th century and died in the 20th.

2.2. The current study

The previous studies focused on the rhythms of instrumental music. In fact, one study indicated they were studying instrumental music because they believed it a foregone conclusion that vocal music would reflect the rhythms of speech, saying: “[i]f music is based on words, and words have different rhythmic properties in the languages under study, then it would be no surprise if musical rhythm reflected linguistic rhythm.” [6]

The current study modified the nPVI measure and applied this modified system to *vocal* music. A large database of musical information was created in David Huron's *Humdrum Toolkit*, an open source software package designed to assist with music research. The database consisted of the melodic, rhythmic, phrasing, and textual information from over six hundred 19th century French and German art songs by nineteen composers, as shown in Table 2.

French		German	
Bizet	(24)	Brahms	(52)
Chausson	(20)	Franz	(61)
David	(34)	Schubert	(59)
Debussy	(29)	Schumann	(36)
Duparc	(14)	Strauss	(33)
Fauré	(37)	Wolf	(82)
Gounod	(29)		
Lalo	(15)		
Massé	(20)		
Massenet	(29)		
Reber	(30)		
Reyer	(8)		
Saint-Saëns	(7)		
Total	(296)	Total	(324)

Table 2. Composers and number of songs represented in the study.

Unlike the strict selection criteria of Patel and Daniele, the primary selection criteria in this study were that the composers had to be native speakers of the language, setting a text originally written in that language by a poet who was a native speaker of the language.²

Because of the different repertory and large size of the database, it was necessary to make some modifications to the method of calculating the nPVI. Rhythmic events such as rests and grace notes turned out to have a significant impact on the nPVI value. These were rhythmic events that were either avoided or modified in the previous studies; all of the previous studies discarded thematic incipits containing grace notes, and Patel and Daniele's study ignored the duration of rests, while Huron and Ollen added the duration of the rest to the previous duration.

The previous studies used as source material short melodic incipits, and as such did not have to include phrasing in their calculations. This study included full

² Stylistic pieces such as *Chinoise*, or any other piece deliberately trying to evoke some “other,” were excluded. The cutoff date for inclusion was 1900.

songs as source material, requiring the nPVI calculations to be modified to calculate the nPVI at the level of the musical phrase.

Calculating the nPVI at the level of the musical phrase coincides with linguistic research on isochrony. There is evidence that perceived isochrony is confined to relatively short speech fragments; that is, the duration of the perceived isochrony has a tendency to change from phrase to phrase and not remain constant over a longer utterance. In fact, linguistic isochrony seems to conform to the principles of the *intonational phrase*, defined as “the unit at which the cognition, physics, syntax, phonetics, and phonology of speech converge.” [9]

Because the nPVI measure is normalized, two phrases can exhibit the exact same nPVI but use different units of isochrony; this accounts for varying rates of speech across (and within) speakers. As a result, it is of limited use to calculate the nPVI of multiple phrases without regard for phrase boundaries, as each phrase will likely be using a different unit of isochrony. It is more fruitful and relevant to calculate the nPVI value of individual phrases.

When applying the nPVI to music, it follows that in order to accurately represent durational variability of rhythmic events, the nPVI measure should be applied to the closest musical correlate to the intonation phrase level of speech, which is the musical phrase.

Therefore, the nPVI measure was modified to calculate the *pnPVI*, or the phrase-nPVI, for individual musical phrases. To calculate the overall pnPVI for larger units, such as entire songs, the pnPVI values for each phrase were weighted based on the number of rhythmic events in the phrase and the weighted results were averaged.

3. RESULTS

3.1. Cross-language results

Figure 1 illustrates the overall weighted average pnPVI value for the French and German songs in the study. The marker represents the weighted overall average pnPVI of each language; the bars represent the range of the overall weighted pnPVI values for composers in that language. While the weighted overall average pnPVI is lower for French than for German (47.9 and 48.1, respectively), this is not a significant result when compared to either the average nPVI values calculated by Grabe and Low [3] of 43.5 and 59.7 or the unexpected nPVI averages found by Huron and Ollen [4] of 43.7 for French and 42.0 for German.

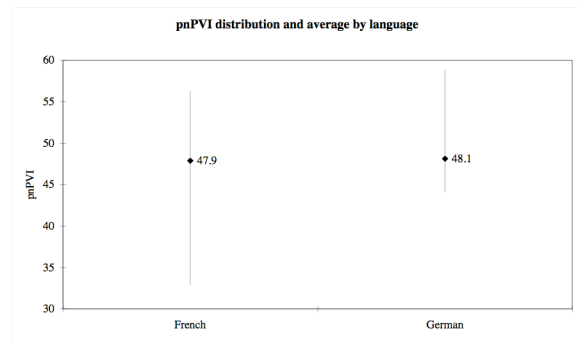


Figure 1. Cross-language average pnPVI .

This result indicates that the hypothesis regarding the rhythmic characteristics of stress-timed and syllable-timed languages and their effect on the rhythmic characteristics of vocal music does not hold for the current set of data. What is most significant about this result is that the hypothesis has failed in a sample of a genre of music where the relationship between language and music has traditionally been assumed to be a very strong one.

3.2. Cross-composer results

Figure 2 shows the weighted average pnPVI value for songs by individual composers and also illustrates the distribution range of weighted average pnPVIs for songs for each individual composer.³ For example, the weighted average pnPVI value for Bizet songs ranges from a minimum of 21.6 to a maximum of 76.8, with a weighted average of 49.7.

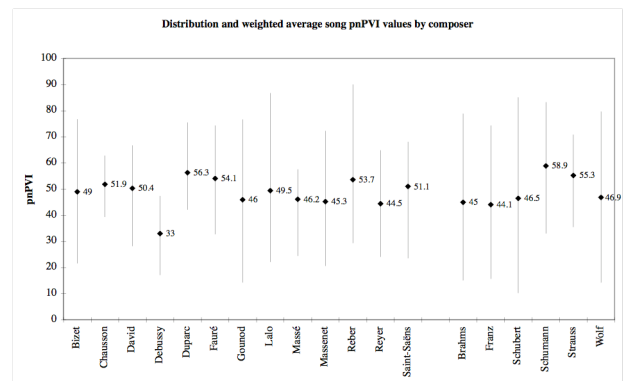


Figure 2. Distribution and weighted average song pnPVI values by composer

Worthy of mention from this graph are the low weighted average pnPVI and low distribution range for the Debussy songs, as well as the narrow distribution range (but relatively high weighted average pnPVI) for Chausson and the relatively high weighted average pnPVI and distribution range for Duparc, Schumann and Strauss. However, with the exception of Debussy, there is no clear trend for the individual pnPVI values to be

³ These averages are calculated by weighting the pnPVI of each song, which itself was calculated by weighting the phrases.

lower for French composers than for German composers.

Therefore, the hypothesis that the syllable-timed French language would influence French composers to write musical melodies with a lower amount of durational variability than that exhibited by music influenced by the stress-timed German language has so far not held for a cross-language study, and, with the possible exception of Debussy, has not held robustly for individual composers.

3.3. pnPVI as a function of time

Figure 3 illustrates the weighted pnPVI of the French songs in the study plotted against the year of their composition. The graph shows the weighted average for the French songs by the decade of composition; the plot for 1840 represents all songs composed between 1840-1849, the plot for 1850 represents all songs composed between 1850-1859, and so on up to 1900, which only represents songs written in the year 1900. The vertical bars represent the range in which the individual songs' weighted pnPVI fell, and the marker and value represent the weighted overall average for songs composed in that decade.

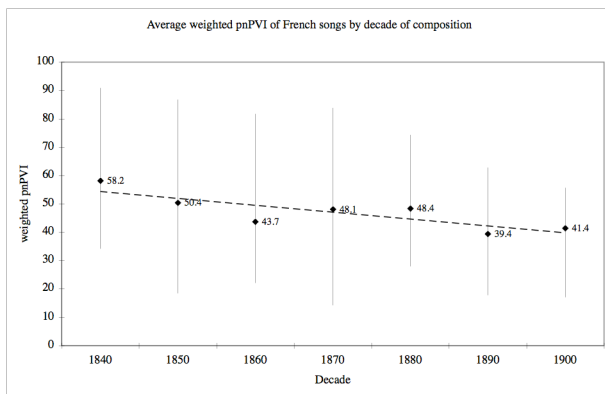


Figure 3. Average weighted pnPVI of French songs by decade of composition

The dotted trend line illustrates the best-fitting regression line (weighted pnPVI = $[-2.4179 \times \text{date of composition}] + 56.757$). The regression of year of composition against average pnPVI is significant ($r^2 = .6817$, $df = 1$, $p = .02$). The chart clearly shows a general trend of declination in average pnPVI value over the time period 1840-1900 for the French songs in this study.

Figure 4 illustrates a similar graph for the German songs in the study. Again, the graph shows the weighted average for the German songs by the decade of composition, the vertical bars represent the range in which the individual songs' weighted pnPVI fell, and the marker and value represent the weighted overall average for songs composed in that decade.

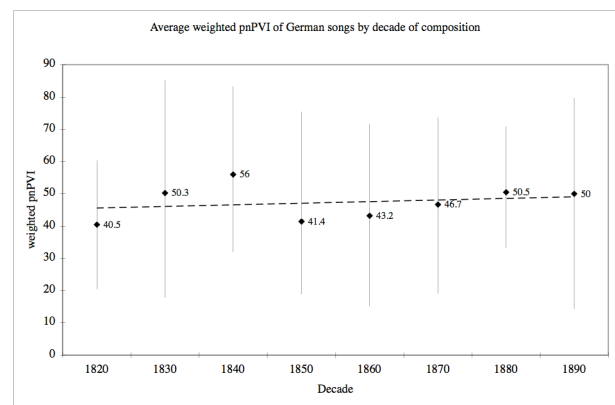


Figure 4. Average weighted pnPVI of German songs by decade of composition

The dotted trend line illustrates the best-fitting regression line (weighted pnPVI = $[0.4929 \times \text{date of composition}] + 45.107$). There is a slight correlation between the decade of composition and an increase in the average weighted pnPVI value. However, the regression of year of composition against average pnPVI is not significant ($r^2 = .0509$, $df = 1$, $p < .6$).

Charting French song pnPVI and German song pnPVI together over time, as in Figure 5, illustrates the trend in each language in relation to the other. While the overall weighted average pnPVI for the data in this study do not support the original hypothesis, for the decade of 1890, the last decade for which there is data for both languages, the pnPVI averages for the two languages (French, 39.4; German 50.0) do support the original hypothesis in a statistically significant way. It is also interesting to note that the trendlines suggest a convergence of French and German pnPVI values in the time period from 1860-1870, and the pnPVI averages for both the 1860s and the 1870s roughly indicate this convergence and subsequent divergence.

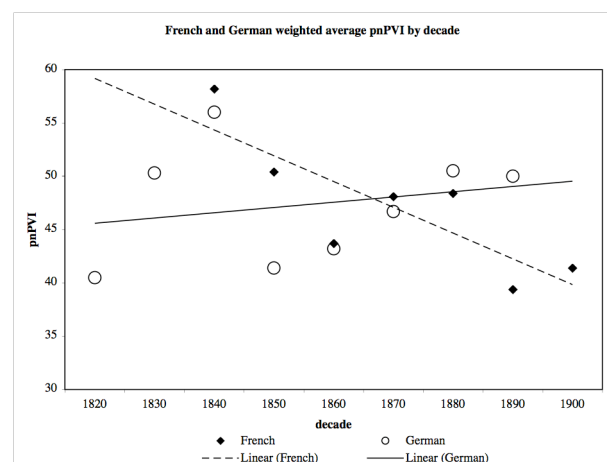


Figure 5. Cross-language pnPVI plot over time.

Figure 6 shows the pnPVI trend of individual French composers based on the repertoire encoded in the current study. This chart illustrates that the overall decline in the pnPVI of French song demonstrated in Figures 3 and 5 is

not the result of data for one or two composers influencing the results. Only two French composers in the study, Massé and Chausson, had an overall trend of increase in pnPVI; the other ten French composers all exhibit a trend of declination in the pnPVI average of their works as the 19th century progressed.⁴ Some composers, such as Gounod and Saint-Saëns, exhibit a sharp overall decline in nPVI value throughout the 19th century; others like Fauré and Reyer show a less pronounced trend of declination. Thus the overall decline in nPVI of French songs is not the result of the influence of any one individual composer on the data, but rather seems to be a general trend among French composers.

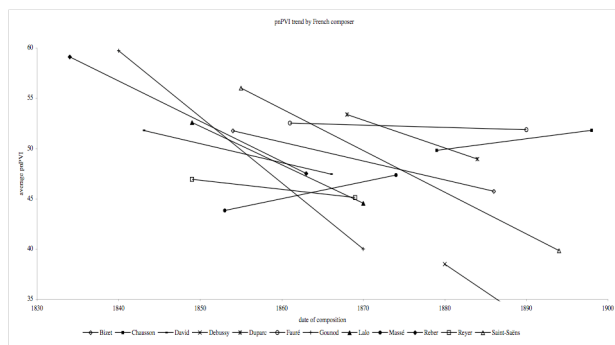


Figure 6. pnPVI results for individual French composers.

4. CONCLUSIONS

The primary hypothesis, demonstrated in previous studies focusing on instrumental music, was that French-influenced vocal music would have a lower nPVI than German-influenced vocal music, reflecting the influence of the rhythmic properties of the spoken language, specifically the greater degree of durational variability present in spoken German than in spoken French. The hypothesis has not held up in a simple comparison between two large groups of songs; the difference between the weighted average pnPVI for all French and German songs was not statistically significant, and with few exceptions there was no clear trend for the weighted overall average pnPVI of individual composers.

However, when pnPVI is plotted as a function of time, there is a clear declination trend for the average pnPVI value of French-influenced music and a slight rising trend for the average pnPVI of German-influenced music throughout the 19th century. These trends are borne out by both the plot of the overall average pnPVI for each language and by the data for the individual composers. It is unlikely that these trends are the result of changes in the spoken language over that time period; therefore, there must have been stylistic changes in French and German song that caused the nPVI to rise or decline. In combination, the two trends result in the difference between the average nPVI values

for each decade approaching and reaching statistical significance by the end of the century. Moreover, these final values match closely the linguistic nPVI results of Grabe and Low [3].

5. REFERENCES

- [1] Beckman, Mary. "Evidence for speech rhythms across languages." In Y. Tohkura, E. Vatikiotis-Bateson, and Y. Sagisaka, eds., *Speech Perception, Production, and Linguistic Structure*. Tokyo: Ohmsha, 1992: 457-63.
- [2] Fox, Anthony. *Prosodic Features and Prosodic Structure: The Phonology of Suprasegmentals*. Oxford: Oxford University Press, 2002.
- [3] Grabe, Esther, and Ee Ling Low. "Durational variability in speech and the rhythm class hypothesis." In C. Gussenhoven & N. Warner, Eds., *Papers in Laboratory Phonology 7*. Cambridge: Cambridge University Press, 2002.
- [4] Huron, David, and Joy Ollen. "Agogic contrast in French and English themes: Further support for Patel and Daniele (2003)." *Music Perception* Vol. 21 No. 2 (Winter 2003): 267-271.
- [5] Laver, John. *Principles of Phonetics*. Cambridge: Cambridge University Press, 1994.
- [6] Patel, Aniruddh D., and Joseph R. Daniele. "An empirical comparison of rhythm in language and music." *Cognition* 87 (2003): B35-B45.
- [7] Ramus, F., Nespor, M. and Mehler, J. "Correlates of linguistic rhythm in the speech signal." *Cognition* 73 (1999): 265-292.
- [8] Toro, J., Trobalón, J., and Sebastián-Gálles, N. "Effects of backward speech and speaker variability in language discrimination by rats," *Journal of Experimental Psychology: Animal Behavior Processes* Vol. 31 No. 1 (2005): 95-100.
- [9] Wennerstrom, Ann. *The Music of Everyday Speech: Prosody and Discourse Analysis*. New York: Oxford University Press, 2001.

⁴ Debussy's trendline goes off the scale of the chart for readability purposes; the trendline ends in 1898 with a value of 28.6.