A Comparison of Rhythm in English Dialects and Music

Rebecca W. McGowan
Wellesley College

Andrea G. Levitt
Wellesley College & Haskins Laboratories

Informal observations have often been made that a country’s language is reflected in its instrumental music. Limited research exists studying similarities between the rhythmic characteristics of French music and language on the one hand and British music and language on the other. Our research compares the rhythmic characteristics of the music and English dialects of the Shetland Islands in Scotland, County Donegal in Ireland, and the state of Kentucky, examining spontaneous speech and unscored musical recordings from the same people. We found that rhythmic characteristics are correlated in the speech and music in each dialect area.

Received February 11, 2010, accepted August 7, 2010.

Key words: speech rhythm, musical rhythm, dialects, nPVI, English

Several recent, empirical studies have investigated similarities between rhythmic characteristics of music and of language in various regions using innovative tools, notably the normalized Pairwise Variability Index or nPVI (Low, Grabe, & Nolan, 2000). The nPVI measures the degree of contrast in the durations of neighboring elements (e.g., vowels or notes). Patel and Daniele (2003) and Patel, Iversen, and Rosenberg (2006) compared British and French speech and music and found that British speech and music featured more durational contrast (a higher nPVI) than did French speech and music. Huron and Ollen (2003) further included music of multiple European nationalities, although they did not measure speech.

The present study expands the previous work by considering dialects of English rather than different languages. In addition, we studied spontaneous speech rather than read speech since significant prosodic differences have been found between the two (e.g., Deterding, 2001; Silverman, Blauw, Spitz, & Pitrelli, 1992). Finally, we studied the production of unscored music as opposed to musical notation, as had been done in the previous studies.

Regions

We examined the speech and music of Kentucky in the United States, Donegal in Ireland, and the Shetland Islands in Scotland, three regions with very distinct dialects and musical styles because of influences unique to each, and because the music is learned and played by ear and without a score. However, the dialects of Shetland, Donegal, and Kentucky are related, with Donegal and Shetland showing Scottish influence, and Kentucky in turn showing Scots-Irish influence (Cooke, 1986; Montgomery, n.d.; Wells, 1982). Musically, the fiddle is an important instrument in each region and a kind of dance melody found in them, the reel, seems to have a common origin in Scotland (Cooke, 1986; Doherty, 1999; Jabbour, 1996).

Dialects

Previous research has shown that the dialects of English differ significantly in their prosodic characteristics, including nPVI. Ferragne and Pellegrino (2004a, 2004b) and White and Mattys (2007) have found significant nPVI differences in passages read by speakers of several dialects of the British Isles. Although few researchers have studied the prosodic characteristics of the dialects of Kentucky, Donegal, and Shetland, some prosodic differences have been found among broader regions: the southern United States, Ulster, and Scotland.

Donegal (Ireland), a variety of Ulster (Northern Irish) English. More uniformity in vowel durations in both Scottish and Ulster dialects than in other English dialects is suggested by the Scottish Vowel Length Rule (SVLR): depending on context, some vowels are not reduced and others are lengthened (Wells, 1982). Furthermore, Ulster English exhibits pitch-peak delay (Grabe, 2002), which White and Mattys (2007) hypothesized results in a leveling out of the contrast in duration in stressed and post-stress syllables. Indeed, Ferragne and Pellegrino (2004b) found that Scottish dialects as well as Ulster dialect (with
are simply less extreme “dotted” rhythms, which players
second in a 4:3, 5:3, or sometimes 2:1 ratio. (These ratios
Appalachia, the first note in a pair being longer than the
between note durations is not as large in Shetland as in
the appropriate “lift” for dancing. However, the contrast
ied from one eighth note to the next in order to provide
Cooke (1986) observed that note durations are also var-
rhythms within the framework of the 4/4 time signature.
Although its vowel lengths are also influenced by the SVLRR, Shetland dialect also fea-
tures a “Scandinavian” syllable structure more promi-
nently (van Leyden, 2004), in which long vowels are
followed by short consonants and short vowels by long
consonants (Catford, 1957, as cited in van Leyden, 2004).
Since this pattern does not reduce the contrast between
long and short vowels, we might expect more variability
in vowel length in Shetland than in other dialects. Indeed,
White and Mattys (2007) found that Shetland dialect
had a significantly higher nPVI score than other dialects
that, like Ulster English, exhibit pitch-peak delay. We
thus expected Shetland dialect to have a higher nPVI
value than Donegal English.
Kentucky (USA). Turning now to the United States,
southern speech is characterized by longer vowels than
northern speech (Wetzell, 2000; Jacewicz, Fox, & Sal-
mons, 2007). In southern speech, stressed vowels can be
lengthened and diphthongized (Feagin, 1991, as cited in
Wetzell, 2000), while unstressed syllables may not be,
resulting in a wider range of vowel durations. Indeed,
Wells (1982) characterized the southern drawl as “involv-
ing relatively greater length in stressed, accented syllables
as compared to unstressed” (p. 529). As part of a transi-
tion zone between southern and midland dialects (Wells,
1982), Kentucky speech likely shares some southern fea-
tures. We thus hypothesized that Kentucky speech would
also have a higher nPVI value than Donegal speech due
to differential lengthening of vowels in Kentucky.

Music

Although the reel, an up-tempo dance melody in 4/4
time, appears to have originated in mainland Scotland
(Cooke, 1986; Doherty, 1999; Jabbour, 1996), each of the
three regions we studied has developed its own specific
rhythms within the framework of the 4/4 time signature.
Musicologist Burman-Hall (1975) noted that musicians
from southern Appalachia were likely to use dotted rhythms where a pair of eighth notes would be played as
a 3:1 ratio rather than evenly (a 1:1 ratio). In Shetland,
Cooke (1986) observed that note durations are also var-
ied from one eighth note to the next in order to provide
the appropriate “lift” for dancing. However, the contrast
between note durations is not as large in Shetland as in
Appalachia, the first note in a pair being longer than the
second in a 4:3, 5:3, or sometimes 2:1 ratio. (These ratios
are simply less extreme “dotted” rhythms, which players
make use of but which are not found notated to such a
precise degree.) Further, Cooke remarked that Irish
music also features long-short note pairs. However, these
contrasts may be smaller still according to the observa-
tion of one musician: “the rhythm tends to become a
very evenly-spaced, equal distribution of quavers” (Ó
Riada, 1982, p. 53). Indeed, Mac Aoidh (1994) men-
tioned that the uneven rhythms characteristic of Scottish
music were leveled out when the music was adopted in
Donegal. Thus we may find the least amount of dura-
tional contrast in Donegal music and the most in
Kentucky.

Method

Source of Material

Speech and music phrases were drawn from archived
field recordings of fiddle players made between 1949
and 1960 (see Appendix A for details). Recordings of
both the music and the speech of the same individuals
were available, and they predated the music revival that
began just after this time, which “undoubtedly has con-
tributed to the undermining of regional identities”
(Vallely, 1999, p. 310). The archived recordings often
included, in addition to fiddle playing, interviews and
storytelling by the musicians, which permitted us to ana-
lyze speech and music spoken and played by the same
person. We included only those musicians who were
recorded speaking at least twenty phrases as defined
below. The speech and music of three fiddle players from
each region were chosen for this study, except in the
Shetland region in which music was only available for
two of the three subjects.

Phrase Selection

Following Grabe, Post, and Watson’s (1999) methodology for analyzing spontaneous speech, we calculated the
nPVI of intonation phrases that were “produced fluently,
without any obvious hesitations, or filled pauses” (p.
1202). We additionally excluded every utterance shorter
than four syllables, containing the quoted words of
another speaker or foreign language, or not ending in
phrase-final intonation, which was determined both by
looking at the F0 contour and listening to the phrase. All
speakers had a minimum of 20 phrases, and we chose 30
phrases at random from those speakers with more than
30. Thus, the speech corpus included 20-30 phrases spo-
ken by three speakers from each of the three regions.
Five reels were chosen at random from each musician. Most reels consist of two main parts of eight bars that
A Comparison of Rhythm

We chose as phrases these two-bar "proposition and answer" segments (McCullough, 1977). From each tune, we extracted four phrases: the first and last phrases of the first and second main parts of the tune, for a total of 20 phrases per player. We chose to analyze the first and last phrases because they are typically distinct, giving a more representative sample of the rhythmic character of each tune. Further, any characteristics emphasized by choice of phrases would be emphasized in all regions due to the structural similarity of the music. An example of a reel, notated, is shown in Figure 1, and the four phrases used in the analysis are marked.

Segmentation

In order to calculate the nPVI, we marked the boundaries of vowels and notes and measured their durations in Praat (Boersma & Weenink, 2009) while viewing a waveform and wide-band spectrogram with a 5 ms window and a frequency resolution of 86 Hz, and using interactive playback.

We marked boundaries of vowels using vowel onset and offset criteria from Peterson and Lehiste (1960), supplementing their guidelines when placing difficult boundaries for /l/, /s/, /w/, /j/, and /h/ by using perceptual cues in combination with rapid changes in formants or energy visible on the spectrogram. We followed Patel et al. (2006) and segmented individual vowels rather than vocalic intervals. In addition, syllabic consonants were treated as vowels as in Ferragne and Pellegrino (2004a).

Each note we considered to begin with its onset and end with the onset of the next note. There were never pauses between notes. Since they are not considered to be notes themselves, ornaments (including slides) were considered to be a part of the note they preceded rather than taken as separate notes or excluded entirely because their function is to articulate the notes being played (Gainza, Lawlor, & Coyle, 2004).

We considered a note to begin with its onset and end with the onset of the next note. There were never pauses between notes. Since they are not considered to be notes themselves, ornaments (including slides) were considered to be a part of the note they preceded rather than taken as separate notes or excluded entirely because their function is to articulate the notes being played (Gainza, Lawlor, & Coyle, 2004). Previous studies such as Patel et al. (2006) measured notations in musical themes that were chosen not to contain ornaments such as grace notes.

Samples of music as well as speech may be heard online at the Digital Library of Appalachia (http://www.ca-dla.org/index.php), the Alan Lomax Archive (http://www.lomaxarchive.com/index.html), and the School of Scottish Studies (http://www.pearl.arts.ed.ac.uk/). Recordings that we analyzed are available at the first two sites, but the actual recordings we used for the Shetland samples are not available online.

1To test if this choice of phrases would influence the results dramatically, we measured all phrases in one tune chosen randomly from each region. The average nPVI value for each region's tune was similar to the values obtained from just the first and last phrases. (Donegal: mean nPVI of all phrases = 24, mean nPVI of first and last phrases = 23; Kentucky: all = 48, first and last = 54; Shetland: all = 26, first and last = 28).

2Samples of music as well as speech may be heard online at the Digital Library of Appalachia (http://www.ca-dla.org/index.php), the Alan Lomax Archive (http://www.lomaxarchive.com/index.html), and the School of Scottish Studies (http://www.pearl.arts.ed.ac.uk/). Recordings that we analyzed are available at the first two sites, but the actual recordings we used for the Shetland samples are not available online.
Future studies might compare the effects of including and excluding such features from the analysis.

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

(1)

The nPVI is a measurement of how much durational contrast exists between neighboring elements and is therefore sensitive to the patterning of the durations. For example, a sentence with long-long-short-short vowels has a lower nPVI score than one with long-short-long-short vowels. The nPVI is calculated using Equation 1, where \( m \) is the number of vowels or musical notes and \( d_k \) is the duration of the kth vowel or musical note.

The nPVI computes the difference between successive vowel durations, which is then normalized by dividing by the mean duration of the pair of vowels. Thus, the nPVI takes \( m \) vowel durations and computes a sequence of \( m-1 \) “contrastiveness scores.” The mean of the scores is then computed and multiplied by 100 to give the nPVI (Low, Grabe, & Nolan, 2000).

Results

General characteristics of speech and music phrases are given in Table 1 and were compared using one-way analyses of variance in SPSS. Speech phrases were not significantly different in duration and number of syllables, but were in speech rate, \( F(2, 246) = 8.61, p < .001 \), with the fastest speakers from Donegal. This difference has no effect on the nPVI, however, since the nPVI normalizes for rate. Musical phrases differed in duration, \( F(2, 157) = 61.45, p < .001 \), and number of notes, \( F(2, 157) = 114.25, p < .001 \), but the means patterned differently from the nPVI results. We did not compare a notes-per-second rate for music since tempo is typically defined in beats per minute.

Speech and music nPVI values are shown in Figure 2. We conducted Kruskal-Wallis tests on both the linguistic and musical nPVI data in order to determine whether there were differences among regions.

In speech, there was a significant between-groups (region) effect on nPVI, \( H(2) = 7.26, p = .027 \). Mann-Whitney \( U \) tests were used for posthoc comparisons

![Figure 2](image_url)

**FIGURE 2.** Speech and music nPVI for each region. (A) Speech nPVI for each region. Donegal mean nPVI = 54.21 (SE = 1.70), Kentucky mean nPVI = 61.77 (SE = 1.90), Shetland mean nPVI = 55.34 (SE = 2.23). (B) Music nPVI for each region. Donegal mean nPVI = 30.05 (SE = 1.37), Kentucky mean nPVI = 46.73 (SE = 1.84), Shetland mean nPVI = 32.24 (SE = 1.75).
A Comparison of Rhythm

between Kentucky and Donegal and between Kentucky and Shetland, and a Bonferroni correction was applied (a .025 level is required to report significance). Kentucky and Donegal were significantly different, \( U = 2691, p = .013 \). The comparison between Shetland and Kentucky approached significance, \( U = 2518, p = .028 \).

Additionally, the differences in music nPVI among regions were highly significant, \( H(2) = 45.07, p < .001 \). Using the same tools as above, we made posthoc comparisons between Kentucky and Donegal, and Kentucky and Shetland, and found highly significant differences between Kentucky and both Shetland, \( U = 482, p < .001 \), and Donegal, \( U = 625, p < .001 \).

**Discussion**

This study adds to a small body of empirical work showing that characteristic linguistic features may be found in the corresponding musical styles. While previous studies have compared different languages and their related music, this study finds that similar relationships may be found in the musical styles and the corresponding dialects of a single language.

Kentucky speech had a significantly higher nPVI score than Donegal, which is consistent with our hypothesis, and can likely be explained by the greater lengthening of stressed vowels than of unstressed vowels in Kentucky speech. Additionally, if Donegal speech shares pitch-peak delay with Ulster English, it may feature a leveling of differences in duration between stressed and unstressed syllables (White & Mattys, 2007), further distancing Donegal from Kentucky in terms of nPVI. We also predicted a higher nPVI for Shetland speech compared to Donegal, but there was no significant difference, although the means were in the predicted direction. The difference between Kentucky and Shetland did not quite reach significance, although the mean for Kentucky, arguably the dialect with the most variable rhythm, was higher than that for Shetland.

We also found differences in musical nPVI values: Kentucky had a significantly higher nPVI value than either Donegal or Shetland. These results are consistent with our hypothesis that Kentucky music would have the highest nPVI value due to the dotted rhythms that Burman-Hall (1975) noted. We note that the differences in nPVI are in the same direction for speech and music, supporting our hypothesis that speech and music would share rhythmic characteristics. Unlike Patel & Daniele (2003), our musical nPVI differences are more dramatic than those for speech. We speculate that the repetitive nature of the musical form of the reel serves to emphasize rhythmic patterns as compared to spontaneous speech.

Our results support findings that dialects of English have different rhythmic characteristics and expand these findings to include spontaneous speech and dialects not previously studied. Turning to music, we have captured some rhythmic differences among the musical styles of three regions that had previously only been observed and described impressionistically. The success of the nPVI at capturing these differences suggests that this measure may be applied in the future to further study of the rhythmic characteristics of music.

**Author Note**

We wish to thank the American Folklife Center, Berea College Appalachian Sound Archives, and the School of Scottish Studies Archives at the University of Edinburgh for granting permission to use their materials. In particular, we would like to thank Cathlin Macaulay at the School of Scottish Studies Archives and Harry Rice at the Department of Special Collections and Archives at Berea College for their help in obtaining materials remotely. This research was conducted as part of a senior thesis and was supported by the Jerome A. Schiff Fund.

Correspondence concerning this article should be addressed to Andrea G. Levitt, French Department, Wellesley College, 106 Central St., Wellesley, MA 02481. E-MAIL: alevitt@wellesley.edu

**Appendix A**

**Sources of Speech Recordings**

**Donegal**

*Mickey Doherty*


*Neil Boyle*


*John Doherty*


**Kentucky**

*Jim Bowles*


1959, Berea College Appalachian Sound Archives. Hutchins Library, Department of Special Collections and Archives, Berea College. Berea, KY 40404.
Sources of Music Recordings

Donegal

Mickey Doherty

Neil Boyle

Kentucky

Jim Bowles

Isham Monday
1959, Berea College Appalachian Sound Archives. Hutchins Library, Department of Special Collections and Archives, Berea College. Berea, KY 40404.

Shetland

John Stickle
1955, School of Scottish Studies Archives. School of Literatures, Languages and Cultures. University of Edinburgh. 27 George Square, Edinburgh EH8 9LD, Scotland.

Fraser Hughson
1955, School of Scottish Studies Archives. School of Literatures, Languages and Cultures. University of Edinburgh. 27 George Square, Edinburgh EH8 9LD, Scotland.

References


