David Huron

Conrad Grebel College
University of Waterloo[1]
Waterloo, Ontario N2L 3G6
Canada
(519) 885-0220, ext. 247
dhuron@watserv1.uwaterloo.ca

The Melodic Arch in Western Folksongs

Many ethnomusicologists have noted that melodic passages tend to exhibit an arch shape where the overall pitch contour rises and then falls over the course of a phrase or an entire melody. In some melodic phrases, this tendency appears to be obvious. However, in any arbitrarily selected phrase, there may be no clear evidence in support of this general claim. For example, many phrases or melodies have pitch contours which show a predominantly ascending or predominantly descending tendency.

When proposing the existence of a general musical feature, such as the melodic arch, it is important to avoid unintentional bias when selecting illustrative passages. Relying on analyses of isolated works rightly raises our suspicions, since the author may have unwittingly selected only those musical examples that exhibit the presumed feature. Counter-examples may fail to be addressed in the discussion.

Large databases of musical scores provide an important resource that allows musicologists to test hypotheses of a general nature by using appropriate sampling methods. The goal of statistical sampling is to provide an equal basis for all instances of a phenomenon to be tallied or analyzed. This is especially desirable when examining every instance is prohibitive or impossible. Large samples

[1] On leave in 1996-97 at the Center for Computer Assisted Research in the Humanities, Braun Music Center, Stanford University, Stanford, CA 94305-3076.

provide especially good opportunities for characterizing both the breadth of variability and the degree of "typicality" in a designated repertory.

The *Essen Folksong Collection* (Schaffrath, 1995) provides a very large sample of (mostly) European folksongs encoded from diverse sources. Each encoding of the 6,251 folksongs in the *Essen Folksong Collection* includes pitch and duration information, meter signatures, bar lines, rests, and explicit phrase markings. The presence of phrasing information in this database provides a unique opportunity to test the prevalence of the so-called melodic arch.

In order to test the arch-contour hypothesis, five separate investigations were carried out using the *Humdrum Toolkit* analysis software (Huron, 1994). The first three investigations pertain to possible arch contours within musical phrases. The fourth investigation focuses on possible arch contours across entire folk melodies. The final investigation examines possible hierarchical patterns of nested arch contours. To anticipate our conclusions, the results support the general hypothesis that a disproportionate number of musical phrases and melodies tend to exhibit an arch-shaped pitch contour—at least in the case of Western folksongs.

Creating Aggregate Phrase Contours

In the first test, all phrases were extracted from the Schaffrath database and sorted according to the number of untied notes in the phrase. Tied notes were treated as single notes and rests were ignored. Of the 36,075 phrases in the database, the phrase lengths range from just one note (one instance) to 30 notes (one instance). The most common phrase length was found to be eight notes (8,532 instances). The median phrase length is also eight notes. Half of all phrases are between seven and nine (untied) notes in length, and three-quarters of all phrases are between six and ten notes in length.

For all phrases, the pitches were translated to numerical values that represent the semitone pitch-distance from Middle C (C4). Thus, C4 was represented as `0` semitones, A4 was represented by the value `9`, C5 by the value `12`, and so forth. Pitches below Middle C were represented by negative integers.

Using this numerical representation, all phrases of a given length (in notes) were averaged together. That is, the semitone heights of all first notes in the phrases were averaged together. Then the semitone heights for all second notes were averaged together. The process was repeated for each subsequent note position within the phrase.

The twelve graphs shown in Figure 1 plot the results of averaging for phrases from 5 to 16 notes in length. Each graph in Figure 1 is accompanied by a number indicating the total number of phrases participating in the calculation. In general, the greater the number of participating phrases, the more likely that the mean pitch heights are good estimates of the actual pitch heights for phrases of that length.

By way of example, the first graph in Figure 1 shows the average pitch contour for phrases consisting of five notes. The initial plotted value indicates that the first notes in five-note phrases have an average pitch height of 8.45 semitones above Middle C. The second notes in all five-note phrases in the database show an average pitch height of 8.78 semitones above Middle C, and so forth. In general, the graphs in Figure 1 show clear inverted "U" shapes which are consistent with the melodic arch concept.

Note that a consistent anomaly appears for phrases longer than 11 notes. In these longer phrases a central dip is evident in the graphs. The cause of this dip is not known. However, its presence suggests a midpoint between two sub-phrases, each showing its own arch shape.

Two possible objections might be raised concerning the graphs in Figure 1. In the first instance, we might ask what proportion of actual phrases exhibit an inverted-U contour. If the majority of phrases exhibit a roughly horizontal or "hovering" contour, then only a few phrases need exhibit an arch contour in order for the average contour to show an upward bulge. In short, it is possible that the graphs in Figure 1 illustrate pitch contours that are characteristic of only a minority of phrases.

A second possible objection is that the arch shapes may be an artifact of combinations of other more complex contours. Suppose, for example, that all phrases conform to one of two predominant melodic shapes. One shape might consist of a hovering contour that then tends to descend toward the end of the phrase. Another shape might consist of a rising tendency at the beginning of the phrase, followed by a hovering pitch that is sustained to the end of the phrase.

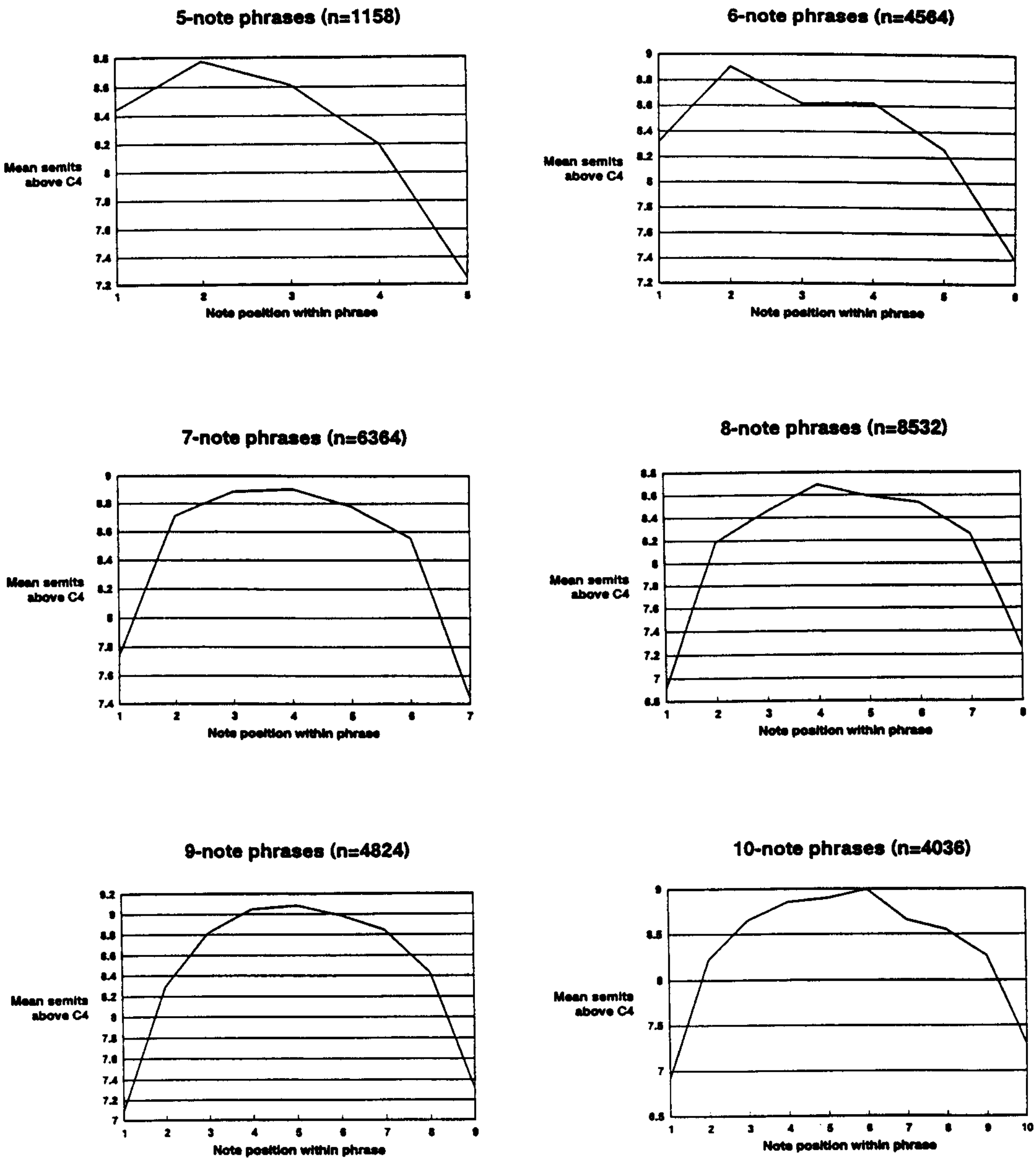


Figure 1 (continues on next page).

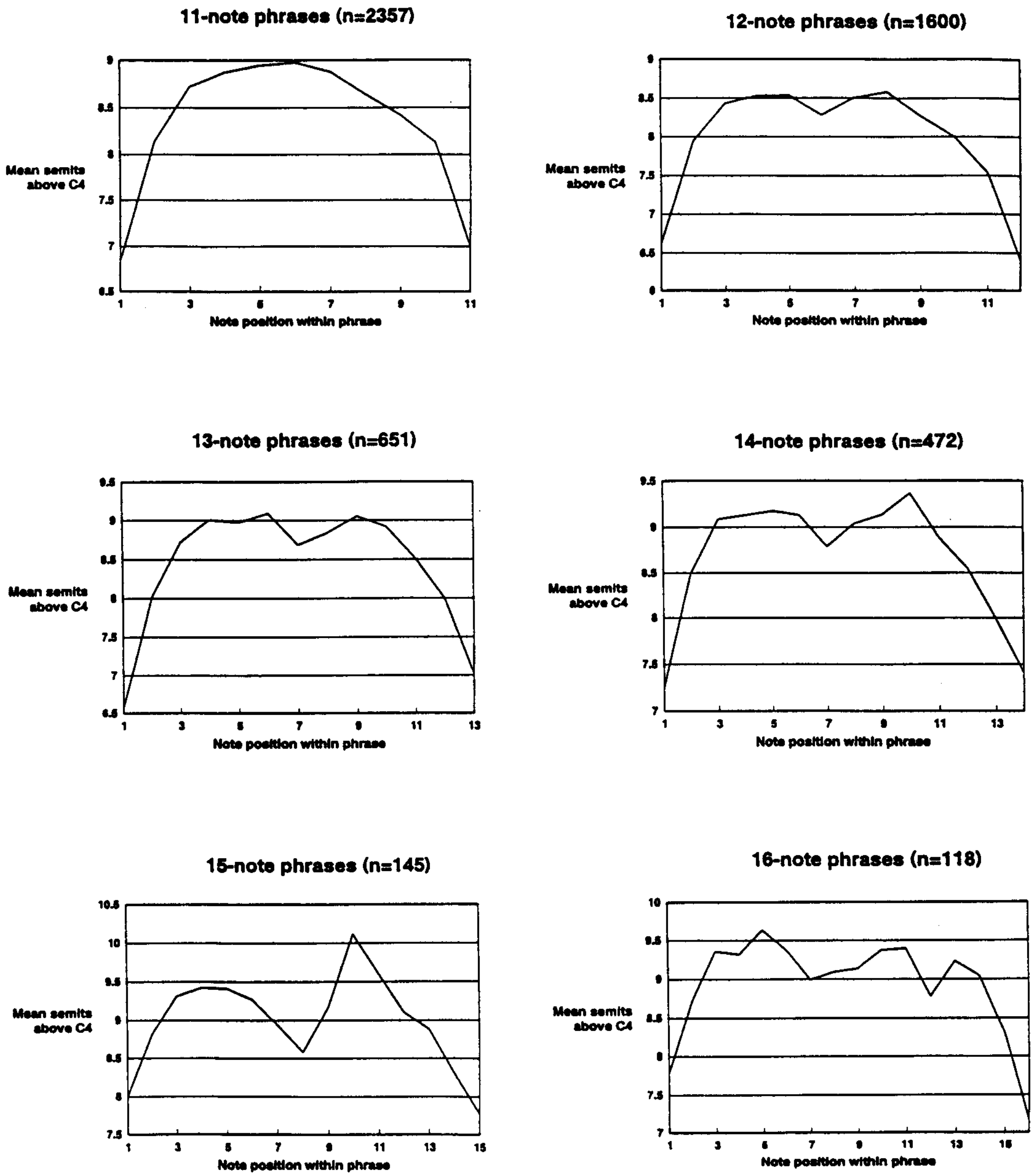


Figure 1. Graphs showing the average pitch contours for all phrases encoded in the *Essen Folksong Collection* (Schaffrath, 1995). Each graph plots the average pitch height (in semitones above Middle C) for successive notes within the phrase. Graphs pertain to phrases of different lengths. The value n indicates the number of phrases participating in the calculation of each graph. Tied notes were treated as single notes; rests were ignored. The figure shows a predominance of convex contours, which is consistent with the concept of the melodic arch.

Averaging together the contours from these two types of phrases will result in an arch shape that fails to characterize accurately the contour of any of the original phrases.

Analysis (Level 1): Individual Phrase Contours

In order to address these two hypothetical confounds, a second series of test measurements that focused on the pitch contours for individual phrases was made. For each phrase in the *Essen Folksong Collection*, the first and last pitches were determined and all the remaining pitches in the phrase were averaged together. Using these three values, all phrases were classified as belonging to one of nine contour categories.

A phrase was deemed to have an *ascending* contour if the final pitch was higher than the average mid-phrase pitch, which in turn was higher than the first pitch in the phrase. Any phrase displaying the reverse relationship was deemed to have a *descending* contour.

If both the first and last pitches in the phrase were higher than the average mid-phrase pitch, then the contour was deemed to be *concave;* the inverse relationship was deemed to be a *convex* contour.

Four types of "dog-leg" contours were also distinguished. If the first pitch and average mid-phrase pitches were equivalent, then the contour was deemed *horizontal-ascending* if the final pitch was higher, or *horizontal-descending* if the final pitch was lower. If the final pitch and average mid-phrase pitches were equivalent, then the contour was deemed *ascending-horizontal* if the first pitch was lower, or *descending-horizontal* if the first pitch was higher. Finally, if the first, last, and average mid-phrase pitch values were equivalent, then the contour for the phrase was classified as *horizontal.*

By way of illustration, Example 1 shows the Danish tune "Kejseren havde de Dottre tre." The first phrase consists of nine notes. Both the first and last notes of this phrase are A4 (*i.e.*, the A nine semitones above Middle C). The average of the seven mid-phrase pitches is (9 + 9 + 9 + 17 + 17 + 17 + 14)/7 or 13.1 semitones above Middle C (roughly approximating to the pitch C♯5. The first pitch, average mid-phrase pitch, and last pitch produce an up-down contour (9

→13.1 → 9) and so the phrase shape is designated "convex." All five phrases in Example 1 have been classified using the same method.

Example 1. The Danish folksong "Kejseren havde de Dottre tre" from the *Essen Folksong Collection.*

Given these contour definitions, each of the 36,075 phrases in the database was classified according to gross shape. If the arch hypothesis is correct, then we would predict a predominance of convex contour types. The classification results are shown in Table 1.

CONTOUR TYPE	NUMBER OF PHRASES	PERCENTAGE
ascending	6,983	19.4%
descending	10,376	28.8%
concave	3,496	9.7%
convex	13,926	38.6%
horizontal-ascending	181	0.5%
horizontal-descending	439	1.2%
ascending-horizontal	307	0.9%
descending-horizontal	174	0.5%
horizontal	193	0.5%
TOTAL	36,075	100.0%

Table 1. Classification of phrases.[2]

[2] An alternative classification method might divide each phrase into three segments of equal length and average the pitches in each segment. When this method is used, the results (shown in the continuation of this footnote on p. 10) differ little from those in Table 1.

On the basis of our pitch-contour taxonomy, 38.6% of the phrases were found to be consistent with a convex or arch-shaped contour. By contrast, only 9.7% of the phrases were consistent with the reverse (concave) contour. Using a chi-squared [χ^2] statistical analysis, the difference in the relative frequencies of occurrence for the concave and convex contours was found to be highly significant (χ^2=6244.1; df=1; p<<0.000001).[3]

Of additional note is the tendency for dog-leg contours to favor patterns similar to the arch. Specifically, dog-leg contours tend to favor lower initial pitches and lower final pitches. Table 1 shows 307 instances of *ascending-horizontal* versus 174 instances of *descending-horizontal* contours; similarly, there are 439 instances of *horizontal-descending* versus 181 instances of *horizontal-ascending* contours. Once again, even in the dog-leg contours, first and last pitches remain more likely to be lower than other pitches in the phrase. Both of these comparisons are statistically significant at better than p<0.000001.

As a final observation, Table 1 shows that the second and third most common contour-types are the descending pitch contour (28.8%) and the ascending pitch contour (19.4%).

Analysis (Level 2): Combinations of Phrases

The relative prevalence of ascending and descending phrase contours raises the possibility that melodic arches may tend to span two or more phrases. That is, it

CONTOUR TYPE	NUMBER OF PHRASES	PERCENTAGE
ascending	6,873	19.1%
descending	9,195	25.4%
concave	4,844	13.4%
convex	12,568	34.8%
horizontal-ascending	516	1.4%
horizontal-descending	778	2.2%
ascending-horizontal	563	1.6%
descending-horizontal	527	1.5%
horizontal	211	0.6%
TOTAL	36,075	100.0%

[3] The x^2 statistical test is explained in the appendix to this article.

may be the case that ascending phrases tend to combine with ensuing descending phrases to form a convex pitch contour.

In order to study such possible phrase combinations, a first-order analysis of phrase-type successions was carried out. The melodic arch hypothesis would predict, for example, that descending phrases would tend to be preceded by ascending phrases. Table 2 computes the types of successive phrase-pairs where all consequent phrases show a descending contour. For example, the first table entry shows that 2,612 cases were found where an ascending-contour phrase preceded a descending-contour phrase.

If successive phrase contours were independent of each other, then one would expect the probability of a given succession to be the same as the percentages given in Table 1. That is, since ascending contours represent 19.4% of all phrases, then one might expect that roughly 19.4% of the phrases preceding a descending contour would be ascending contours.

The actual percentages are shown in the third column of Table 2. The differences between zero-order probabilities and first-order probabilities are given as percentage changes. Plus signs indicate that the antecedent phrase is more common than would be expected in a random arrangement of successive phrases. Conversely, minus signs indicate that the antecedent phrase type is less common than would be expected by chance. The final two columns in Table 2 show the results of individual χ^2 analyses.

ANTECEDENT CONTOUR	NUMBER OF INSTANCES	PERCENTAGE OF INSTANCES	PERCENTAGE CHANGE	χ^2 [CHI-SQUARED]	P( ) [PROBABILITY]
ascending	2,612	27.3%	(+41.2%)	246.25	0.000001
descending	2,271	23.8%	(-17.4%)	67.26	0.000001
concave	1,117	11.7%	(+20.6%)	35.13	0.000001
convex	3,203	33.5%	(-13.2%)	48.15	0.000001
horizontal-ascending	61	0.6%	(+27.2%)	3.52	0.06063 N.S.
horizontal-descending	82	0.9%	(-29.5%)	10.04	0.00153
ascending-horizontal	113	1.2%	(+38.9%)	12.17	0.00048
descending-horizontal	48	0.5%	(+4.1%)	0.08	0.77730 N.S.
horizontal	52	0.5%	(+1.7%)	0.01	0.92034 N.S.
TOTAL	9,559	100.0%			

Table 2. Distribution of phrase-types preceding descending phrases.

Table 2 shows six statistically significant results. The presence of a descending-contour phrase tends to increase the likelihood of an antecedent ascending-contour phrase, or an antecedent concave contour, or an ascending-horizontal phrase. The presence of a descending-contour phrase also tends to suppress an antecedent descending-contour phrase, or an antecedent convex contour, or a horizontal-descending phrase. In summary, a descending-contour phrase tends to follow a phrase that ends on a higher pitch. This finding is consistent with the melodic arch hypothesis.

Analysis (Level 3): Contour Balance in Successive Phrases

A possible objection to the preceding analysis is that since melodies tend to maintain a fixed pitch range or tessitura, it follows that increases in pitch height ought to increase the probability of a subsequent decrease in pitch. That is, the high pitches that tend to terminate an ascending phrase contour ought to be followed by contours that tend to return the melody to the central pitch region.

In order to address this possible confound, a further "counterbalanced" analysis was carried out. The concept of the melodic arch predicts that ascending phrases ought to be followed by descending phrases. However, it does not necessarily predict that descending phrases ought to be followed by ascending phrases. Since the "tessitura return" tendency would be expected to apply to both low pitches as well as high pitches, there is no reason why the number of descending-ascending phrase-pairs should differ from the number of ascending-descending phrase-pairs. In general, the arch hypothesis predicts that "what goes up must come down," but does not dictate that "what goes down must come up."

Consequently, we might predict the following phrase-order asymmetries in accordance with the melodic arch hypothesis:

1. The number of ascending → descending phrase-pairs should be greater than the number of descending → ascending phrase-pairs.
2. The number of concave → descending phrase-pairs should be greater than the number of descending → concave phrase-pairs.
3. The number of convex → descending phrase-pairs should be greater than the number of descending → convex phrase-pairs.

4. The number of horizontal-ascending → descending phrase-pairs should be greater than the number of descending → horizontal- ascending phrase-pairs.
5. The number of ascending-horizontal → descending phrase-pairs should be greater than the number of descending → ascending- horizontal phrase-pairs.

No predictions would be made regarding descending → descending, horizontal-descending, descending-horizontal, and horizontal phrase contours. The premise underlying all of these predictions is that antecedent phrases that end with ascending pitch contours ought to precipitate a consequent fall in pitch.

Table 3 tabulates the various pairs of possible counterbalanced phrase orders. The prime phrase orders (ending with descending phrases) are shown first, whereas the retrograde phrase orders (beginning with descending phrases) are shown below. Percentage differences in the number of instances are identified in the right-most column of Table 3. Plus signs indicate that the differences are in the direction predicted by the asymmetrical hypotheses listed above.

With one exception, all of the results are skewed in the direction consistent with the melodic arch hypothesis. Chi-squared analyses show that only the first two of the above hypotheses are statistically significant. That is, the number of ascending → descending phrase-pairs is greater than the number of descending → ascending phrase-pairs (χ^2=74.78; df=1; p<0.000001); and the number of concave → descending phrase-pairs is greater than the number of descending → concave phrase-pairs (χ^2=21.82; df=1; p<0.000001).

PRIME ORDER	No.
ascending → descending	2,612
descending → descending	2,271
concave → descending	1,117
convex → descending	3,203
horizontal-ascending → descending	61
horizontal-descending → descending	82
ascending-horizontal → descending	113
descending-horizontal → descending	48
horizontal → descending	52
TOTAL	9,559

RETROGRADE ORDER	NO.	% DIFFERENCE
descending → ascending	1,522	+6.6%
descending → descending	2,271	[no pred.]
descending → concave	684	+2.4%
descending → convex	2,563	+1.4%
descending → horizontal-ascending	44	0.0%
descending → horizontal-descending	84	[no pred.]
descending → ascending-horizontal	77	+0.1%
descending → descending-horizontal	42	[no pred.]
descending → horizontal	53	[no pred.]
TOTAL	7,340	

Table 3. Phrase-pair incidences: (a) in prime order and (b) in retrograde order.

Analysis (Level 4): Melodies as Arches

Given the above results, we might finally consider whether whole melodies may be regarded as consistent with some sort of arch contour. Although some of the simplifications we have used in the foregoing analyses may be appropriate to musical phrases, it is less likely that these simplifications are suitable for characterizing entire melodies. Indeed, the complexity of the measurement issues raises cautionary flags.

One approach might seek to determine whether the middle notes of a melody tend to be higher in pitch than the initial and final notes. However, since we have already established that phrases have a tendency to exhibit arch shapes, and since melodies typically begin with the beginning of a phrase and end with the end of a phrase, we would necessarily expect the initial and final pitches in a work to be lower than some arbitrarily selected group of mid-melody pitches. Similarly, since we have established that phrases have a tendency to form groups of ascending-descending pairs, it follows that initial and final phrases of a melody would be more likely to display ascending and descending contour types respectively.

An alternative approach might average all pitches within individual phrases so that each phrase is represented by a single average value. We might then determine whether the average phrase heights themselves tend to conform to an ascending-descending profile over the course of the melody.

Given the crude quality of such measurements, this approach would seem to be inauspicious and unlikely to reveal any arch shape. However, this proves not to be the case. The results in Table 4 show that over 40 per cent of average pitches for successive phrases do indeed show a convex contour.

CONTOUR TYPE	NUMBER OF WORKS	PERCENTAGE
ascending	1,706	27.3%
descending	1,020	16.3%
concave	805	12.9%
convex	2,621	41.9%
horizontal-ascending	18	0.3%
horizontal-descending	13	0.2%
ascending-horizontal	23	0.4%
descending-horizontal	18	0.3%
horizontal	27	0.4%
TOTAL	6,251	100.0%

Table 4. Average pitches of successive phrases.

These results are illustrated graphically in Figure 2, which shows average melodic contours for all of the works in the database. Each graph in Figure 2 plots the average pitch height (in semitones above Middle C) for successive phrases within the melodies.

Two general trends are evident in these graphs. First, there is a marked tendency for the average pitch of successive phrases to increase over the course of the melody. Second, for melodies having four or more phrases, there is a clear tendency for most mid-melody phrases to have a higher average pitch than the initial and final phrases. Once again, this latter feature is consistent with the melodic arch hypothesis.

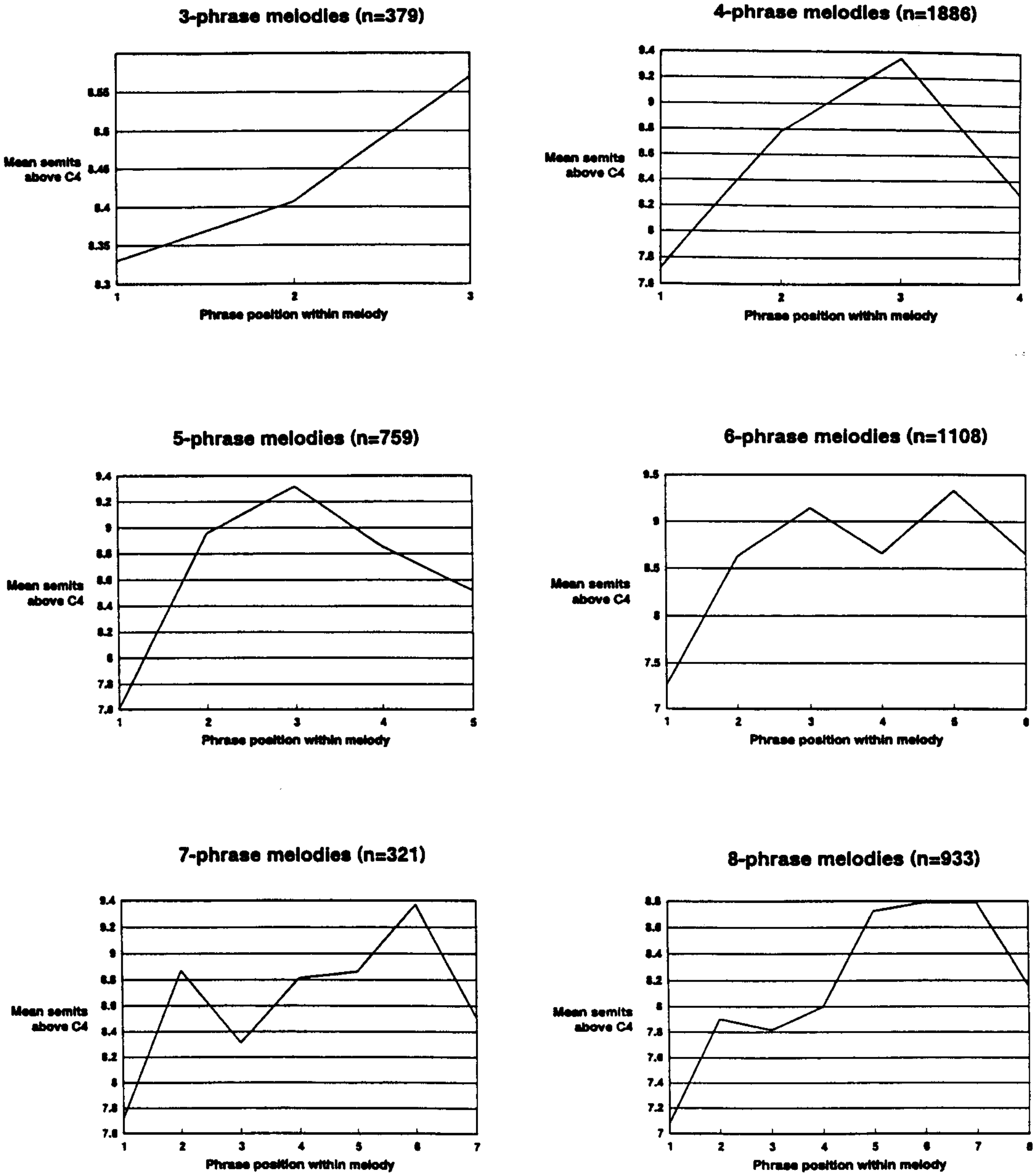


Figure 2. Graphs showing the average pitch contours for all melodies encoded in the *Essen Folksong Collection* (Schaffrath, 1995). Each graph plots the average pitch height (in semitones above Middle C) for successive phrases within the melodies. Graphs pertain to melodies containing different numbers of phrases. The value indicates the number of melodies participating in the calculation of each graph. Tied notes were treated as single notes. The figure shows the evidence of convex contours which are consistent with the concept of the melodic arch.

Analysis (Level 5): Are Arches Hierarchical?

The preceding results suggest that arch-shaped pitch contours appear to occur at two or three different temporal levels. As we have seen, both phrases and entire melodies exhibit arch tendencies. Moreover, long phrases appear to tend to break into sub-phrases, with each sub-phrase showing its own arch tendency (recall Figure 1, later panels).

In addition, successive pairs of phrases may cooperate in the creation of arch contours. These observations raise the question of whether arch contours tend to be nested in particular melodies. That is, are melodies that exhibit an overall arch contour more likely to contain phrases that also display arch contours? Or conversely, do melodies that exhibit an overall arch contour tend to suppress the presence of arches at the phrase level?

In order to address these questions a correlational study was carried out. At least two predictions arise from the arch-hierarchy hypothesis:

(1) We would predict a positive association between the proportion of phrases in a melody that are classified as convex and whether the overall melody itself is classified as convex.

(2) We would predict a positive association between the proportion of phrase-pairs that are ascending-descending and whether the overall melody itself is classified as convex.

In order to test these hypotheses, each melody in the database was assigned a *convexity value* of either `1` (convex) or `0` (non-convex) according to the criteria outlined above. In addition, the proportion of phrases within each melody that are classified as convex was determined.

For example, a four-phrase melody that contains three convex phrases was assigned the value 0.75. If arch-shaped phrase contours tend to be nested in arch-shaped melodic contours, then we would predict a significant positive correlation between the whole-melody convexity and the proportion of phrases that are convex. For 6,251 melodies, the Pearson's coefficient of correlation was measured to be +0.015 (non-significant).[4]

In order to test the second hypothesis, the *proportion of phrase-pairs* within a melody that display ascending-descending contours was determined. Proportions were calculated based on the maximum possible number of ascending-descending phrase-pairs, given the number of phrases in the melody. For example, only two ascending-descending phrase-pairs are possible in both four-phrase melodies and five-phrase melodies. Hence a five-phrase melody exhibiting two ascending-descending phrase-pairs would be assigned a value of 1.00.

If arch-shaped phrase-pairs tend to be nested in arch-shaped melodies, then we would predict a significant positive correlation between the whole-melody convexity and the proportion of phrase-pairs that are ascending-descending. For the entire *Essen Folksong Collection* the correlation value was measured to be -0.017 (non-significant).

As a final test, an *aggregate index of phrase-convexity* was assembled. This index identified the proportion of all phrases in a melody that are either convex or that participate in an ascending-descending phrase-pair. For example, a four-phrase melody having one ascending-descending phrase-pair and one convex phrase would have a value of 0.75 (*i.e.*, three of four phrases participate in some type of arch contour). The purpose of this aggregate index was to better characterize the tendency of phrases within a melody to exhibit convex contours. Calculating Pearson's coefficient of correlation resulted in a value of +0.001 (non-significant).

In general, the results of these correlational tests are consistent with the view that there is no particular relationship between melodies that show an overall arch contour and arch-shaped phrases or phrase-pairs. On the basis of these observations we would reject both the claim that arches tend to be nested and the claim that arches tend to occur only at one temporal level in a melody.

[4] Pearson's coefficient of correlation is used to characterize the degree of similarity between two sets of numbers. Correlation values range between +1 (exactly similar) and –1 (exactly contrary). Correlation values near zero indicate that the two sets of numbers have no similarity in their fluctuations.

Conclusion

In summary, the above investigations provide considerable evidence consistent with the notion of a melodic arch. In the first instance, averaging pitches for equal-length phrases consistently produced contours exhibiting an inverted-U shape. In the second instance, nearly 40 per cent of phrases were found to have initial and final pitches that are lower than the average mid-phrase pitch.

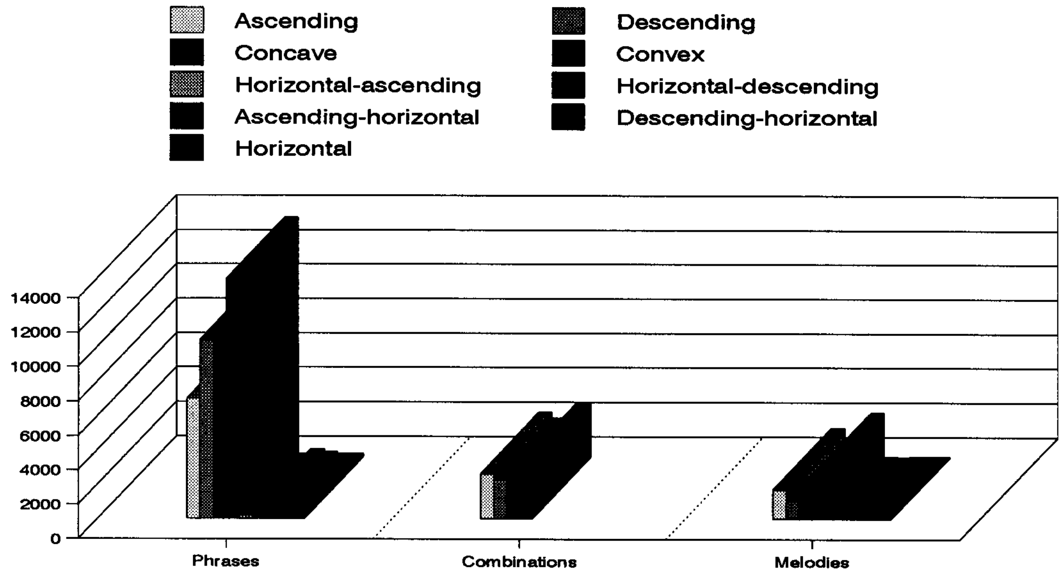
Of nine simple contour types, a convex shape was found to be the most common. In the third instance, while ascending and descending pitch contours are relatively common, there is a significant tendency for ascending and descending phrases to be linked together in pairs. Moreover, these pairings cannot be attributed merely to the tendency to maintain a given tessitura.

What goes up is likely to come down, but what goes down is less likely to come back up. In the fourth instance, reducing each phrase to an average pitch produced contours over entire melodies that exhibit an inverted-U shape. Over 40 percent of melodies were found to have initial and final phrases that exhibited lower average pitches than mid-melody phrases. Of nine simple contour types, an overall convex shape was found to be the most common. Selected results are compared in Figure 3.

Of course, not all phrases or groups of phrases display an arch shape. Indeed, it is possible that as many as half of the phrases in the melodies studied do not exhibit or participate in a convex pitch contour. Similarly, fewer than half of the melodies display an overall arch shape. Nevertheless, at both the phrase and whole-melody levels, the arch shape remains a disproportionately frequent feature of the melodies examined in this study.

Although arches appear to occur at different temporal levels, there is no evidence that arch contours tend to be nested. That is, melodies that show an overall arch contour are not more likely to contain phrases that are also arch-shaped. Nor is it the case that arches at the whole-melody level tend to suppress the presence of arch-shaped phrases within the melody.

All of the above conclusions remain valid only for vocal phrases found in Western folksongs. It is not known whether these features are more or less prevalent in instrumental melodies, or in music from other genres or cultures. A further limitation of this study was the failure to differentiate notes according to their perceptual importance. No attempt was made to distinguish such tones or to focus on the contours implied by the most salient or structurally important notes. The possible influence of such differences, and differences arising from other repertories, will need to await further investigation.



Ascending	6983	2612	1706
Descending	10376	2271	1020
Concave	3496	1117	805
Convex	13926	3203	2621
Horizontal-ascending	181		18
Horizontal-descending	439		13
Ascending-horizontal	307		23
Descending-horizontal	174		18
Horizontal	193		27

Figure 3. Comparison of nine contours in single phrases, successive phrase-pairs ("combinations"), and overall melodies.

References

Huron, David. *UNIX Tools for Music Research; The Humdrum Toolkit Reference Manual.* Menlo Park, CA: Center for Computer Assisted Research in the Humanities, 1994.

Schaffrath, Helmut. *The Essen Folksong Collection in the Humdrum Kern Format,* D. Huron (ed.). Menlo Park, CA: Center for Computer Assisted Research in the Humanities, 1995.

Appendix I. The Chi-Squared Test

The goal of statistical analysis is to estimate the probability of truth or falsity for some statement. A statistical analysis typically ends with the calculation of a value called *p* [= probability] which ranges between 0 and 1. A value near 0 means that the probability of the statement's being true is very low; a value near 1 means that the probability is high.

Statisticians have devised innumerable procedures for calculating *p* depending on the type of data, the quality of the data, and the nature of the claim being examined. Each valuative procedure is called a statistical *test*. There are many kinds of statistical tests; these are often identified by the names of their inventors (*e.g.*, Hartley's F-test, the Wilcoxon signed-ranks test, the Kolmogorov-Smirnov test, and the Kaiser-Meyer-Olkin test).

Since the *Humdrum* tools can generate quantitative data by querying musical data, many claims about music can be tested using statistical methods. The kinds of statements that might be evaluated by statistical means are suggested by those that follow:

- Beethoven used a lot of dynamic changes in his music
- Brahms used a lot of hemiolas
- Gershwin avoided syncopated melodies
- Dowland's lute music shows little influence of the instrument

One of the most common statistical tests is the chi-squared test. This test takes its name from the Greek letter χ (pron. KĪ), which is used to compare proportions or ratios. It facilitates the determination of whether the proportion of occurrences of some features in one data sample is significantly greater than or less than the proportion of the same feature in a different data sample.

By way of example, let us test the hypothesis that Brahms uses hemiolas more than his contemporaries do. To set up the test we need first to count the number of hemiolas in (a) a sample of data derived from Brahms's music and (b) a sample of data by other composers who were active at the same time. Let us suppose that the raw data looks like this:

COMPOSER(S)	MEASURES SEARCHED	HEMIOLAS FOUND
Various	5000	23
Brahms	500	9

Proportionally, hemiolas occurred in 1.8% of the data for Brahms but just 0.7% in the other data. Is this a big difference? A chi-squared test will tell us.

Calculating χ^2

This test entails two parts. First we calculate the value of chi-squared. This is done by subtracting the number of expected occurrences from the number of observed occurrences, squaring the result, and dividing it by the number of expected occurrences. Where O is the observed number of occurrences and E is the expected number, the formula is:

$$\chi^2 = \frac{(O-E)^2}{E}$$

If we would normally expect 23 hemiolas in 5000 measures of music of the period, then we would expect about 2.3 hemiolas in the 500 measures of Brahms's music, i.e., that $E = 2.3$. Since the actual number encountered was 9, the numeric substitutions would be as follows:

$$\chi^2 = \frac{(O-E)^2}{E} = \frac{(2.3-9)^2}{2.3} = \frac{(-6.7)^2}{2.3} = \frac{44.89}{2.3} = 19.5565$$

Our chi-squared value is 19.5565.

Calculating p

The value of p, which tells us what the probability of the number of occurrences is, can be determined by referring to a standard table of critical values of χ^2. For one *degree of freedom* (*df*), we find the following values:

df	.500	.200	.100	.050	.025	.020	.010	.001
1	0.455	1.642	2.706	3.841	5.024	5.412	6.635	10.827

The χ^2 value for a p of 0.001 is 10.827. Our χ^2 calculation gives a greater value, so the p value must be less than 0.001. This means that there is less than one chance in 1000 that one would expect to see such a discrepancy between Brahms and his contemporaries.

Other Considerations

What value of p is considered good? The short answer is that it depends on how certain you want to be. In typical research, a relationship is considered *significant* when the value of p is less than 0.05. For more stringent research, a value of less than 0.01 is considered necessary. The value 0.05 is sometimes referred to as the *beta confidence level*, whereas the value 0.01 is referred to as the *alpha confidence level.*

If the value of p achieves the alpha confidence level, this does not mean that the hypothesis is true. It simply means that it is probably true. Similarly, if the value of p fails to achieve even the beta confidence level, this does not mean that the hypothesis is false. It simply means that the evidence in its favor is not strong.

The value of χ^2 is sensitive to the number of observations. The greater the number of observations, the greater the likelihood that the result will be considered significant. If a coin were flipped 6 times and came up heads 4 of them, this incidence would be considered to lie within the realm of chance. But if the coin came up heads on 400 of 600 tosses (preserving the same proportion), the value of p would be less than 0.001.

There are many other aspects to statistical inferences that we have not discussed here. The purpose of this very brief introduction is to provide an example of how statistics may be used to provide additional evidence pertaining to musicological questions. For any given hypothesis, the scholar must also seek converging evidence from a wide variety of sources and points of view.