Testing Models as Predictors of the Rivalry Between Structure and Surface in the Perception of Melodies

Isabel Cecilia Martínez & Favio Shifres
Universidad Nacional de La Plata

Introduction

Reductional Music theories propose a sort of musical listening which involves a mental processing of a series of events belonging to the musical surface, establishing a hierarchy where some of them receive more structural importance, and, in that way, prolonging their "existence". Harmony defines the tonal space in which prolongation develops itself, interacting with the melody which impulses the movement (Salzer and Schachter, 1969). Both the notes which are prolonged as long as time passes by and the relationships of tension and relaxation between events are important features of this interaction. This musical organisation is substantially governed by the voice leading principles derived mainly from the strict counterpoint. Beyond the musical surface an underlying organisation takes place, whose perceptual reality is intended to investigate. The process of abstraction of main events from the complete musical piece, called reductional representation (Mc Adams, 1989), favours the attribution of tonal coherence to a musical piece (Salzer, 1962).

Research has been developed to study this phenomenon from a psychological perspective (Serafine, 1988; Serafine, Glassman & Overbeeke, 1989; E. Bigand, 1990,1994; N. Dibben, 1994; Martinez & Shifres 1999a, 1999b; Shifres & Martinez 1999). They used different experimental paradigms: goodness of fit between a melody and the rendered reduction of its underlying voice leading (henceforth UVL), similarity judgement between melodies with same or different UVL and family categorisation according to resemblance of melodies.

Nevertheless, the study of reductional representation faces methodological difficulties in order to apply an experimental paradigm: given that the events belonging to the underlying hierarchic levels also belong to the musical surface, when an structural event is modified to create appropriate experimental conditions, the surface is simultaneously modified. Thus, it is difficult to specify the level on which the listener’s response is based.

The study of the cognitive reality of the reductional representation should begin by the analysis of the reciprocal influences between the surface and UVL attributes. Our own research (Martínez & Shifres 1999a, 1999b; Shifres & Martinez 1999), was devoted to monitor this relationship using a similarity judgement paradigm. Results revealed that neither the contour hypothesis (according to which two melodies are judged as more similar as higher is the association between their contours) nor the structure hypothesis (which predicts that two melodies will be judged as more similar if they have same UVL) could explain separately the perceptual similarity. Thus, a perceptual rivalry hypothesis was formulated.

This work follows a previous experiment which aimed to test such hypothesis, analysing further melodic components in order to verify previous interpretations of the results, and to differentiate surface components from those of the UVL. It is an epistemological exercise which asks which kind of melodic knowledge does the listener use while judging similarity of melodies. It is assumed that the explanations provided by alternative models -more formalised and accepted than the one previously employed- may be useful to the analysis of both components. It is expected to bring further support to the exploration of the psychological reality of the assumptions of the reductional representation of the tonal hierarchic structure.

The Baseline Experiment section summarises relevant aspects of the cited investigation in order to clarify the actions which followed it. Next section, The Models, synthesises the advantages of four models to analyse melodic attributes. Two of them (Combinatorial Model and Oscillations Model) focus on surface components, studying the note to note level relationships and considering only the particular melodic information of the musical examples. The other two (Tonal Weights and Melodic Anchoring) emphasise structural aspects of the tonal melodies, because they are based on the study of invariants of the tonal system. The Method section describes the analysis of the melodies used in the former experiment in terms of the mentioned four models. In Results section, empirical data from the experiment are interpreted from the point of view of these new analysis. Finally, in the last section, contributions of these models and different paradigms are discussed.
The Baseline Experiment

A previous study (Martínez & Shifres, 1999b) intended to create experimental conditions to test the hypothesis of Rivalry between structure - understood as UVL - and surface - measured according to the note-to-note relationships of the melodic shape. Stimuli consisted on 15 trios of melodies. They were the result of composing for each of the 15 selected tonal melodies one melody with the same UVL and similar surface level and another of different UVL and similar surface according to the concept of melodic diminutions (Schenker, 1979; Forte & Gilbert, 1982) - Figure 1. In order to find a theoretical measure of similarity between surfaces, a number of parametric controls was used while composing. One of them, the Correlational Model of theoretical melodic shape similarity (Shifres & Martínez, 1999), appeared to be the most relevant to study the Rivalry. This model gives a theoretical measure of similarity between melodic shapes as a function of the Pearson product-moment correlation coefficient between the series of numbers which describe interval relationships of successive notes. A number is assigned to each interval between two successive notes, with the absolute value representing the number of semitones of the interval and the sign + or - corresponding to its ascending or descending direction. Thus, the association between series which represent a pair of melodies belonging to a trio, provides a theoretic account of their relative similarity compared to the other two pairs of the trio. The model was used to monitor the composition of the trios of melodies in order to obtain half of them with the highest association between melodies A and C (setting the AC Group) and the rest with the highest correlation between B and C (BC Group). The lowest association always corresponded to the pair AB. As melodies A and B had the same UVL, it was estimated that this treatment would create a rivalry condition between both surface similarity and deeper structural levels similarity.

Figure 1. Procedure followed in the composition of the stimuli. Example No. 8: Chopin, Study Op. 25 No. 5. a) selection of the fragment (Melody A); b) analysis of the underlying voice leading; c) reduction of the underlying voice leading (R1); d) transformation of R1 into R2; e) Reconstruction of a melody (Melody C) from R2; f) Reconstruction of a melody (Melody B) from R1, homologating changes between B and A to the changes between C and A.

146 adults with different levels of musical experience took part in the experiment. The experimental task consisted on listening to the sequence AB - AC (or AC - AB) and to judge which of the two comparison melodies (B or C) was the most similar to A. Besides, subjects had to estimate the level of certainty of the answer using a three point scale (very sure- not so sure- not sure). It was hypothesised that listeners would judge melody B as the most similar given that it has the same UVL, although different degree of association between surfaces would cause confusions in the responses. Thus, responses for melodies belonging to AC Group would be less certain than responses for BC Group melodies. Data for B/C responses and certainty ratings were translated into a single score ranging from 1 (very certain C) to 6 (very certain B), where 3 and below represent "C" and 4 and above represent "B". Thus, the test value was 3.5.

Results confirmed the prediction: (a) subjects always tend to judge melody B (same UVL) as the most similar; and (b) different levels of scores showed that structural and superficial attributes compete, causing different levels of perceptual Rivalry.

When the UVL is modified, inevitably the surface level is modified as well. For that reason, the experimental control of the theoretical rivalry was based in the identification and exhaustive control of the surface attributes, which in spite of all the precautions taken change anyway.

The purpose of the present work is then to test the pertinence of alternative theoretical models to describe those melodic attributes which were modified while modifying the UVL. It is expected that if the models are useful in giving account of a different melodic information they may help in finding a more precise estimation of the real incidence of the UVL in the similarity judgements. So, stimuli used in the previous experiments were analysed according to the following models: Oscillations Model, Combinatorial Model, Tonal Weights and Melodic Anchoring.
The Models

Oscillations Model (Schmuckler, 1999)

Proposed by Schmuckler (1999), it is probably the most simple idea to describe the melodic contour, considering the most superficial level of note-to-note relationships. In order to differentiate this measure from others used in former studies, and at the same time to capture the melodic information avoiding any type of structural component, it was used the simplest version of the model which consists on counting the number of ascents and descents of the melodic contour.

Combinatorial Model (Marvin & Laprade, 1987)

It was proposed by Marvin and Laprade (1987) in order to analyse melodic information about contour relationships (patterns of ascents and descents) between non adjacent notes. These relationships may be represented in a \( n \times n \) matrix, in which \( n \) represents the number of notes of the melody and each entry is \( = 1 \) if it is an ascent or \( = 0 \) if it is a non ascent (being either a descent or a repetition). The overall melodic direction is easily noticed observing the density of 1s or 0s that are placed above the diagonal of the matrix. The CSIM (Quinn, 1999) is a measure of the similarity between contours of the same number of notes that reveals the proportion of similar entries between the matrixes representing both contours. Its value ranges from 0 (if the matrixes have not common entries) to 1 (if all the entries are common to both matrixes). "This model is sensitive to the interval size as well as interval direction, but only the interval size relative to the other intervals within a pitch pattern" (Quinn, 1999, p. 446). The model may give information about the overall direction of the melody. It also focuses the problem of long term attending (which is present in the listening of UVL) without considering the underlying structure.

Thus, both models (a) capture the information of ascents and descents without considering the interval dimension and (b) represent two edges in relation to the temporal focus required (from the note-to-note level -Oscillations Model-, to the level which considers the long term linear connections that may be present in this type of short melodies - Combinatorial Model ).

Tonal Weights (Krumhansl, 1990).

Tonal Weights describe the Tonal Hierarchy (Krumhansl, 1990) providing quantifiable information of the amount of stability of every melodic tone related to a tonal context. As a psychological principle the Tonal Hierarchy explains the abstraction of the relative stability as an invariant, a sort of implicit knowledge which is acquired by acculturation and stored in the long term memory as a schema. It is activated while listening to a musical piece and contributes to generate expectations concerning to the more stable incoming events. These, in turn, facilitate the codifying of each tone of the piece. It is expected that the model may give an alternative explanation to the incidence of structural components (regardless to the voice leading) in the similarity judgements, providing a measure for the overall tonal stability of the melody. Thus, if two melodies have similar level of tonal stability they will be judged as more similar compared to other melodies with very different levels of stability. Besides, if this measure is combined with an estimation of the note duration, then it would provide a situational description of the event hierarchy within the context of the melody. As the stimuli analysed are constrained by the limits of the psychological present, it is assumed that Tonal Weights contributes to the similarity judgement, in spite of appearing to be a static measure.

Model of Melodic Anchoring (Bharucha, 1984)

Melodic Anchoring (Bharucha, 1984a; 1996) is a psychological principle which characterises the listener's implicit knowledge to assign a position of relative stability or instability to each incoming pitch event of a given melody. Through this principle the more unstable tones tend to be assimilated by the more stable tones which form the tonal schema. Two constraints govern the activation of this principle: (a) asymmetry, (the stable tone always follows the unstable tone and not the other way round) and (b) proximity (two successive events may not be more distant than a major second).

Although the model analyses the note-to-note level, it distinguishes between different levels of stability and describes their relationships providing a measure of the structural stability of the melody. Thus, this principle acts as another structural force, being more dynamic than the Tonal Weights but more local than the UVL. So, it is assumed that this model might capture information that listeners would take into account while judging similarity.
Method

Oscillations Analysis

The number of reversals in the direction of the melody was counted. This provided a measure of the tendency of movement. Although the original version of the model does not include repetitions, as in this test three melodies with the same rhythm were listened, it was assumed that repetition would be clearly noticed as a change of direction (in that way _reversal_ is understood as change of direction). Two measurements were obtained:

1. Differences of oscillations (*) of melodies B and C compared to melody A:

\[ \text{RSIM} = \text{rsim}_{AB} - \text{rsim}_{AC} \]

where \( \text{rsim}_{AB} = \frac{1}{2} \sqrt{R_A - R_B} \); and \( R \) is the number of changes of direction.

2. Classification of trios according to the highest \( \text{rsim}_{XY} (***) \):

- **AB Group**: \( \text{rsim}_{AB} < \text{rsim}_{AC} \)
- **CB Group**: \( \text{rsim}_{AB} > \text{rsim}_{AC} \)

![Figura 2. Procedure followed to compare the theoretic similarity of the melodies according to the four models analysed.](image)

Combinatorial Analysis

Each melody was represented with a matrix. The matrices were compared into pairs counting the number of similar entries and dividing it by the total number of entries of the matrix (without the diagonal- see figure 2). These proportions (csim) generated two measures:

1. Difference of proportions (*):

\[ \text{csim}_{AB} - \text{csim}_{AC} = \text{CSIM} \]

This value ranges from 1 to -1. If positive, A and B have the highest theoretic similarity; if negative, this relationship is between A and C.

2. Classification of trios (***) according to the highest theoretic similarity corresponding to AB (AB Group) or to AC (BC Group)
Analysis of the Tonal Weights

On the assumption that changes of the comparison melody do not affect the structural importance of the notes which do not change, there were considered only different notes between the melodies. The procedure was as follows:

1. Given the pair A and B, the different notes between them were assigned the value of the probe tone ratings of Krumhansl’s test [TW] (1990; p.30) -Figure 2.
2. Product of Tonal Weight:

$$\text{TWP} = \text{TW} \times D$$

Where D is the duration of the note, taking the minimal duration of the fragment as unit.
3. Sum of Products of Tonal Weights:

$$\sum \text{TWP} = \sum \text{TWP}_i + \ldots + \text{TWP}_j$$

Where i, j are the different notes between A and B
4. Difference of Tonal Weights

$$\Delta \text{TW} = \sum \text{TWP}[A] + \sum \text{TWP}[B]$$

If the tonal weights are similar, $\Delta \text{TW}$ will be next to 0 - and the perceptual similarity between both melodies will be higher.
5. The same procedure was followed with AC. Then, the following measures were obtained

1. Proportion of differences between tonal weights (*). Given that subjects listened to the sequence AB - AC, it is possible to think that the relative value of $\Delta \text{TW}$ in a pair within the total $\Delta \text{TW}$s of the whole sequence may influence the answer. Thus,

$$\Delta \text{TW\%} = \frac{\Delta \text{TW}_{AB}}{\Delta \text{TW}_{AB} + \Delta \text{TW}_{AC}}$$

The more similar the amounts of tonal weights between B and C, the more the proportion tends to 0.5. In this case tonal weights will not influence the subject's decision for any of the two melodies. As soon as this proportion is closer to 1 - indicating that the difference in tonal weight of AB is higher than the difference of AC - the similarity judgements will tend to C. Thus, it is possible to estimate a correlation between the $\Delta \text{TW\%}$ and the means of perceptual similarity observed in the experiment.

2. Classification of trios (**) according to their $\Delta \text{TW}$. Trios were classified according to the lowest $\Delta \text{TW}$ found between the pairs AB and AC:

- AB Group ($\Delta \text{TW}_{AB} < \Delta \text{TW}_{AC}$), AC Group ($\Delta \text{TW}_{AB} > \Delta \text{TW}_{AC}$).

Analysis of the Melodic Anchoring

According to the model, the expectation strength of resolution is determined not only by the proximity between an unstable tone and a stable one but also by the relative proximity of the two potential anchor tones which are placed at both sides of the unstable tone. There are other factors which contribute to the expectation strength such as metric position, level of activation and components of attention which are mentioned but not formalised in the model. A force vector, sum of all the forces which exert a pull on an event, is proposed. It represents the direction (up or down) and strength of the pull, conveying a psychological force, an expectation of resolution of a current event to the nearest stable tone. A tonal force vector in a particular direction is proportional to the activation of the nearest anchor (stable tone) in that direction and inversely proportional to the distance -measured in semitones- from that anchor. The value of a tonal force vector results from the difference between the values of both potential tonal force vectors (ascendant and descendant).

Melodic Anchoring (Bharucha,1996) was applied to the 15 trios of melodies according to the following procedure:

1. It was analysed the implicit harmony of the melodies applying the principle of parsimony, if necessary, according to the criteria of consonance, coherence of the harmonic chain, harmonic rhythm and strength of the harmonic relationships.
2. A value of the anchoring strength (ANCH) was assigned to every notes. The chord tones were assigned 0 representing the condition of stability.

- The ANCH for the non chord tones was calculated assigning:
  1. a constant - 1 - which represents the category of unstable tone, plus...
  2. the corresponding value of the tonal force vector $tE$ which qualifies the relative instability of the tone, plus...
Where $A$ represents factors of attention (it is consider constant in the current analysis).

$a^+$ and $a^-$ are the levels of activation of the up and down anchor tones. Here, they are considered roughly the same ($a^+ = a^-$). Thus:

$$tE = (1/d^+)-(1/d^-)$$

When the chord changes between the non chord tone and the chord tone, the non chord value is calculated averaging the value of this tone regarding both chords.

It was added 1 to each value if the tone anchored in the predicted direction, 1.375 if it resolved in the opposite direction and 1.75 if it did not anchored.

3. The average of ANCH of all the tones of the melody was obtained. It represents the value of tonal stability of the melody measured in terms of melodic anchoring. The higher the value the lower the stability of the melody. Thus, it was estimated:

1. Difference of anchoring $\Delta$ANCH.

$$\Delta$ANCH = $\frac{1}{2}$ANCH_A - $\frac{1}{2}$ANCH_B

According to this value, the model predicts that: if the value is negative, the most similar melody is C. And if the value is positive, the predicted highest similarity corresponds to B.

2. Classification of trios $\mathbf{(**)}$ according to their $\Delta$ANCH

AB Group ($1/2$ANCH_A - $1/2$ANCH_B $\lt 1/2$ANCH_C),

AC Group ($1/2$ANCH_A - $1/2$ANCH_B $\gt 1/2$ANCH_C).

Thus, each model provided two measurements: $\mathbf{(*)}$ a continuous one and $\mathbf{(**)}$ a categorical one (which allows the stimuli classification in groups according to their theoretic similarity). They may contribute to analyse the concept of rivalry within each category.

**Results**

The continuous measurements obtained were used to run a lineal regression analysis. It showed that none of the models could predict the differences in the similarity judgements.

Predictions emerging from the classification of the models taken in pairs were compared: Anchoring vs. Oscillations; Anchoring vs. Combinatorial Model; Anchoring vs. Tonal Weights; Tonal Weights vs. Oscillations; Tonal Weights vs. Combinatorial Model; and Oscillations vs. Combinatorial Model. According to this comparison the melodies could represent different possibilities of agreement/disagreement between models: Agreement B (when the trio belongs to the AB Group in both models and they agree on predicting the highest similarity of B); Agreement C (the trio belongs to the AC Group in both models and they agree on predicting the highest similarity of C); and Rivalry (when one model predicts the highest similarity of B and the other predicts the highest similarity of C).

It was predicted that the ratings of perceptual similarity for the melodies classified as Rivalry in the comparative analysis of the models would be intermediate between examples in Agreement B and examples in Agreement C (Figure 3).
Figura 3. Values of predicted perceptual similarity for the cases of agreement and disagreement about the theoretic similarity predicted by a pair of models.

A ANOVA 3 (Agreement/Rivalry) X 5 (Comparisons) repeated measures showed significant results for the factor Agreement/Rivalry ($F[2, 143] = 91.224; p < .000$). Thus, Agreement B represented higher perceptual similarity of B, Agreement C represented higher relative tendency toward C, and an intermediate value represented those cases in which the comparison of the two models resulted in alternative predictions (Rivalry) (Figure 4).

However, none of these means is lower than 3.5, showing that subjects always judge B as the most similar. This implies that the models do not capture all the information which is being used in the similarity judgements.

The factor Comparisons was also significant ($F[4, 141] = 24.616; p < .000$), indicating that the combined effect of the two models within the pair presented differences. This means that, as it is observed in the figure, each pair of compared models show different values of perceptual similarity on the agreement for B or C. For example, the perceptual similarity of B is better predicted by the agreement between the Combinatorial Model and the Melodic Anchoring or the Tonal Weights than by the agreement between the Oscillations Model combined with any of the other models.

The most curious result is that the Interaction between both factors also was significant ($F[12, 134] = 11.906; p < .000$) indicating that if a pair of models represents the best combination in order to predict the perceptual similarity of B, it is not the case for the predicted similarity of C and vice versa. Besides, it can be observed in the figure that although the tendency of highest similarity can be predicted by different combinations of two models, the rivalry between them only is well defined by the combination between the Oscillations Model and the Anchoring Model.
Figura 4. Values of observed perceptual similarity for the cases of agreement and disagreement in the predicted theoretic similarity by 5 different pairs of models

### Discussion

This study tested different models of melodic analysis in order to investigate whether listeners capture the attributes which they describe while judging similarities. The aim was to contrast the conclusions of a previous study according to which, under certain constraints, the Underlying Voice Leading was the attribute that listeners took into account when they made similarity judgements, even though if a Rivalry between the melodic contour and the UVL could be inferred. The results of the contrast of the alternative models show that:

a. It also exists a Rivalry between the attributes -captured by the models- and the UVL;

b. In spite of that Rivalry, choices always favoured the melody with the same UVL, being the mean always higher than 3.5.

c. The UVL is an attribute which listeners seem to capture when they judge similarities. The different models do not explain these responses.

d. Therefore, and according to the tendencies observed in these studies, it is considered that the UVL was isolated as an experimental variable.

One of the main methodological difficulties that presents the experimental study of UVL as a plausible explanation of the cognition of musical structure, is that principles that govern their relationships are neither described in parametric terms nor according to categories defined in absolute terms. In Psychology, it is considered that the analysis of similarities and differences of data is a valid methodological tool in order to capture the structure of the underlying structure (non measured) of the objects under observation. Thus, the application of the paradigm of similarity judgements in the Baseline Experiment was pertinent to the purpose of the former study.

Another methodological challenge has to do with the fact that, although the theory exhibits prescriptive terms, the derivation of these principles is an interpretative matter. The current study praised models which analyse the processing of the melodic information in parametric ways. Thus, the contrast with the explanations providing by such approaches with the "heuristics" of the UVL, intended to abduct the later from the rest of the components. The measurements gave account of an important scope of features characterised by the models, from the note-to-note level of the musical surface to the invariants of the musical structure.

Although the Melodic Anchoring principle characterises some aspects of the melody, invoking voice leading principles, there are differences between this theoretical framework and the one provided by the schenkerian theory. Thus:

a. Melodic anchoring is prospective, while melodic diminutions may be either prospective or retrospective - prefix or suffix (Forte & Gilbert, 1982) -.
b. The constraint of temporal proximity is not pertinent to the principles prescribed by the theory of voice leading in which events of the musical surface may be related to structural events which are relatively distant.

c. Unlike Melodic Anchoring, voice leading principles establish hierarchies between tones not only as a result of the relationship chord tone-non chord tone but also as a result of their power to establish a coherent horizontal relationship between the previous and the following tone at a particular level of the hierarchy. Thus, for example, the concept of melodic diminutions of a structural tone may imply arpeggiation and consonant jump as a more superficial status, while from the Melodic Anchoring point of view these notes represent chord tones of the same structural level.

Results indicate that none of these models gave a more parsimonious explanation of the similarity judgements than the one provided by the theory of UVL. Therefore, we can go further in recognising the validity of this theory to bring an explanation of the facts. Nevertheless, it is interesting to observe the way in which the models also rival, evidencing that each of them are giving account of different attributes of the melody.

Considering the similarity judgement as a process of feature-matching that is the result of operations of contrast between collections of features which allow to weigh what is common against what is different (Tversky, 1977), the present analysis show that, in fact, the models provide only partial explanations of such differential features and that there is "something more" in the responses, something that the listener grasps in order to configure the response. According to the characteristics of the stimuli, whose surfaces were carefully monitored, this component is the UVL.

The relationship between Music Theory and Psychology of Music has involved the study of relatively basic concepts such as chords, scales, intervals, etc. Success in this primary relationships between both disciplines has encouraged the development of experiments in order to test the cognitive reality of more complex theoretical constructs (Krumhansl, 1995). This implies both a methodological and an epistemological challenge. In this case, the results obtained encourage the future application of the model of contrast (Tversky, 1977) to induce aspects of the tonal structure at deeper levels.

References


Back to index