



**HAL**  
open science

## Objective Evaluation of Tonal Fitness for Chord Progressions Using the Tonal Interval Space

Marcelo Caetano, María Navarro-Cáceres, Gilberto Bernardes

► **To cite this version:**

Marcelo Caetano, María Navarro-Cáceres, Gilberto Bernardes. Objective Evaluation of Tonal Fitness for Chord Progressions Using the Tonal Interval Space. *Artificial Intelligence in Music, Sound, Art and Design*, 12103, Springer International Publishing, pp.150-164, 2020, Lecture Notes in Computer Science, 10.1007/978-3-030-43859-3\_11 . hal-03473384

**HAL Id: hal-03473384**

**<https://hal.science/hal-03473384v1>**

Submitted on 10 Dec 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Objective Evaluation of Tonal Fitness for Chord Progressions Using the Tonal Interval Space

María Navarro-Cáceres<sup>1</sup>, Marcelo Caetano<sup>2,3</sup>, and Gilberto Bernardes<sup>4</sup>

<sup>1</sup> Department of Computer Sciences, University of Salamanca. Pza de los Caídos, s/n. 37007 Salamanca, Spain. [maria90@usal.es](mailto:maria90@usal.es)

<sup>2</sup> CIRMMT, Schulich School of Music, McGill University, Montreal, Quebec, Canada  
[marcelo.caetano@mcgill.ca](mailto:marcelo.caetano@mcgill.ca)

<sup>3</sup> Aix-Marseille Univ, CNRS, PRISM “Perception, Representation, Image, Sound, Music”, Marseille, France [marcelo.caetano@prism.cnrs.fr](mailto:marcelo.caetano@prism.cnrs.fr)

<sup>4</sup> INESC TEC and University of Porto, Faculty of Engineering, Porto, Portugal  
[gba@fe.up.pt](mailto:gba@fe.up.pt)

**Abstract.** Chord progressions are core elements of Western tonal harmony regulated by multiple theoretical and perceptual principles. Ideally, objective measures to evaluate chord progressions should reflect their *tonal fitness*. In this work, we propose an objective measure of the fitness of a chord progression within the Western tonal context computed in the Tonal Interval Space, where distances capture tonal music principles. The measure considers four parameters, namely tonal pitch distance, consonance, hierarchical tension and voice leading between the chords in the progression. We performed a listening test to perceptually assess the proposed *tonal fitness* measure across different chord progressions, and compared the results with existing related models. The perceptual rating results show that our objective measure improves the estimation of a chord progression’s *tonal fitness* in comparison with existing models.

**Keywords:** Chord Progression, Hierarchical Tension, Tonal Interval Space, Melodic Attraction, Consonance.

## 1 Introduction

Chords are fundamental elements of Western tonal music. The vertical construction of chords and its horizontal motion, known as chord progressions, have been the subject of several theories over the past decades [15, 17, 18]. Most theoretical analyses of chord progressions focus on particular elements among the multidimensional principles regulating chord progressions, such as consonance, musical tension and voice leading. Among these theories, we can highlight those that express tonal pitch relations in topological spaces [8, 17] aiming to capture the sense of proximity between chords in Western tonal music. In the aforementioned theories, chords in a progression are addressed both linearly and hierarchically, thus highlighting the importance between consecutive chords and their function

across multiple hierarchies. In this paper, the objective measure of how well a given chord fits a progression is called “*tonal fitness*”<sup>5</sup>.

Some authors propose to measure the *tonal fitness* of the next chord in a progression using the previous chord as reference, therefore, only considering the linear dimension of a chord progression. Quick *et al.* [13] present a grammar with probabilities based on a previous analysis of Bach pieces. Woolhouse *et al.* [21] propose the pitch attraction model, which evaluates one chord according to the previous one. Callender *et al.* [6] present a generalized space for chord representation to evaluate two consecutive chords. There are proposals which evaluate chords by considering more than a single previous chord in the progression [7, 12, 20]. These works analyze style-specific music corpora to statistically extrapolate long-term features, such as typical movements of tonal functions or common melodic sequences. However, the resulting models do not capture long-term dependencies such as phrase structures. The evaluation of the fitness of the chords within the hierarchical dimension of music structure requires a hierarchical analysis.

To address this limitation, several authors propose models in which the tonal properties of a chord are measured by considering not only on the previous chords but also their hierarchical relationships, typically represented as a tree structure [8]. Bernstein draws a basic relationship between music and Chomsky’s formal grammars [5]. Schenker [17] proposes a hierarchical analysis to reduce the musical surface to tonal functions, which lacks a comprehensive computational formulation. Steedman [19] designs a context-free grammar to model Blues progressions. Rohrmeier [16] presents a grammar to generate structures for tonal music based on hierarchical trees that can capture different tonal hierarchies. Lerdahl’s [8] proposal to measure tonal tension and melodic attraction of a chord progression from linear and hierarchical structures of a musical phrase is one of the most influential models to date. Lerdahl adopts four tonal indicators computed manually from chord progressions: *stability*, *consonance*, *hierarchical tension*, and *voice leading*.

In our work, we depart from the conceptual basis of Lerdahl’s model to computationally measure the *tonal fitness* of a chord within a progression. Tonal pitch indicators inspired by Lerdahl’s model are computed in our work using the Tonal Interval Space (TIS) by Bernardes *et al.* [4], where hierarchical tonal pitch relations are expressed as distances. Use of the discrete Fourier transform (DFT) in the TIS to calculate the distances results in a computationally efficient representation. Besides, this representation automatically defines the regional (or key) space of a given chord progression, in contrast with Lerdahl’s model, which requires manual definition. Therefore, our work results in a more flexible computational analysis framework which also broadens the scope of target applications

---

<sup>5</sup> Lerdahl’s and Farbood’s proposals use the term *tonal tension* to refer to how a chord is hierarchically related to the rest of the chords in a progression. Following this concept, we will use “tonal fitness” to capture how well a chord reflects tonal properties in the context of the chord progression, according to the Western tonal rules. It therefore includes Lerdahl’s concept of tonal tension

and users by excluding the need for prior musical knowledge or annotations on the analyzed musical data.

Furthermore, our work addresses both linear and hierarchical long-term dependencies of chord progressions, which also extend the manually-driven Lerdahl’s proposal with an automatic computational method for representing musical hierarchies from a given chord progression as tree structures. In short, stemming from state-of-the-art models, our approach not only fosters a computationally efficient framework for the analysis of chord progressions’ *tonal fitness*, but also aims to improve the accuracy of tonal indicators by adopting the perceptually-inspired TIS. Ultimately, we believe that our contribution can support and enable new tools for the automatic analysis and generation of hierarchically-aware tonal chord progressions, whose lack of hierarchical structure has been a long identified problem in the field of generative music [10].

To validate our measure, we conducted a listening test with chord progressions from different tonalities. We compare our currently proposed measure with a previous work [11] whose measure only considers the linear dimension of the chord progressions, and with Lerdahl’s [8] measure. Comparison with a previous work allows to analyze how the hierarchical and linear dimensions can influence the subjective ratings in the same representation space, whereas comparison with Lerdahl’s measure aims to assess how the representation space can affect the measurement of the *tonal fitness*.

The remainder of this paper is structured as follows. Section 2.1 describes the TIS and its features. Section 2 explains the theoretical paradigm that influences the formalization of the measure. Section 3 describes the measure to calculate the *tonal fitness* of the chords following linear and hierarchical criteria. Section 4 details the experiment, and Section 5 describes the results obtained and the discussion. Finally, Section 6 presents the conclusions and the future work.

## 2 Theoretical Background About Chord Progressions

### 2.1 Tonal Interval Space

Fourier analysis has been recently used to explore tonal pitch relations [2, 4, 14]. Bernardes *et al.* [4] proposed to calculate the weighted DFT of a chroma vector  $C(n)$  to obtain Tonal Interval Vectors (TIVs). The TIVs can be used to represent different tonal pitch hierarchies, such as pitch classes, intervals, chords, and keys.

One of the most important properties of the TIS is that distances among different pitch configurations represented by their TIV reflect musical attributes. Each TIV is interpreted as vector  $T(k)$  with  $k = 1, \dots, M = 6$ , where  $M$  is the number of components considered from the DFT. The Euclidean distance between  $T_1(k)$  and  $T_2(k)$  is given by

$$\rho(T_1, T_2) = \sqrt{\sum_{k=1}^M |T_1(k) - T_2(k)|^2}. \quad (1)$$

Perceptually, similar vectors result in smaller distances than dissimilar ones. The inner product between  $T_1(k)$  and  $T_2(k)$  is

$$T_1 \cdot T_2 = \sum_{k=1}^M T_1(k) \overline{T_2(k)}, \quad (2)$$

where  $\overline{T_2(k)}$  is the complex conjugate of  $T_2(k)$ . Eq. (2) yields higher values for perceptually similar vectors than dissimilar ones. The angle  $\theta$  between  $T_1$  and  $T_2$  can be calculated from eq. (2) as

$$\theta(T_1, T_2) = \arccos \left( \frac{T_1 \cdot T_2}{\|T_1\| \|T_2\|} \right), \quad (3)$$

where  $\|T_1\|$  denotes the magnitude of  $T_1$  calculated as the distance  $\rho(T_1, T_0)$  between  $T_1$  and the center of the space  $T_0$  using eq. (1). Eq. (3) results in smaller angles to indicate a higher degree of similarity.

Eq. (1) to eq. (3) can be used to compute distances between pitch configurations of the same level (e.g., the distance between two chords) or across levels (e.g., the distance between a chord and a key). For example, the distance between two chords captures their tonal relatedness, the distance between a chord and a key measures the level of membership of the chord to the key, and the distance of a given chord to the three categorical harmonic functions (represented in the space as the triads of the tonic, subdominant and dominant degrees) measures how well the chord fits the categorical harmonic functions. All these properties are used in the proposed measure (Section 3) to capture the suitability of the chord in the context of a progression.

## 2.2 Lerdahl’s Theory to Model the Tonal Fitness of Chord Progressions

Some authors have proposed models of structural dependencies of chord progressions [8, 16, 17]. In particular, Lerdahl [8] proposed a measure  $L(T_i, P)$  that assesses the tonal tension and the melodic attraction of a chord  $T_i$  in a progression  $P$ . Lerdahl’s model explores the role a chord plays within the hierarchical structure of a chord progression by considering four elements, namely, *tonal pitch distance*, *surface dissonance*, *voice leading*, and *hierarchical structure*.

*Tonal pitch distance* captures the proximity of chords using an algebraic representation. Lerdahl proposes a space that measures chord distances using a lookup table, in which non-common tones between chords, key distances, and the interval distance among the chords root are considered [9].

*Surface dissonance* measures the psychoacoustic dissonance of chords in a progression from the interaction among the vertical component notes of each chord. Surface dissonance results from the combination of three factors: “scale degree”, which considers the component scale degrees in a given chord; “inversion”, which accounts for the chord’s bass note and on its subsequent elaboration

as root position or inversion; and “nonharmonic tone”, which inspects the existence of tones outside the chord function.

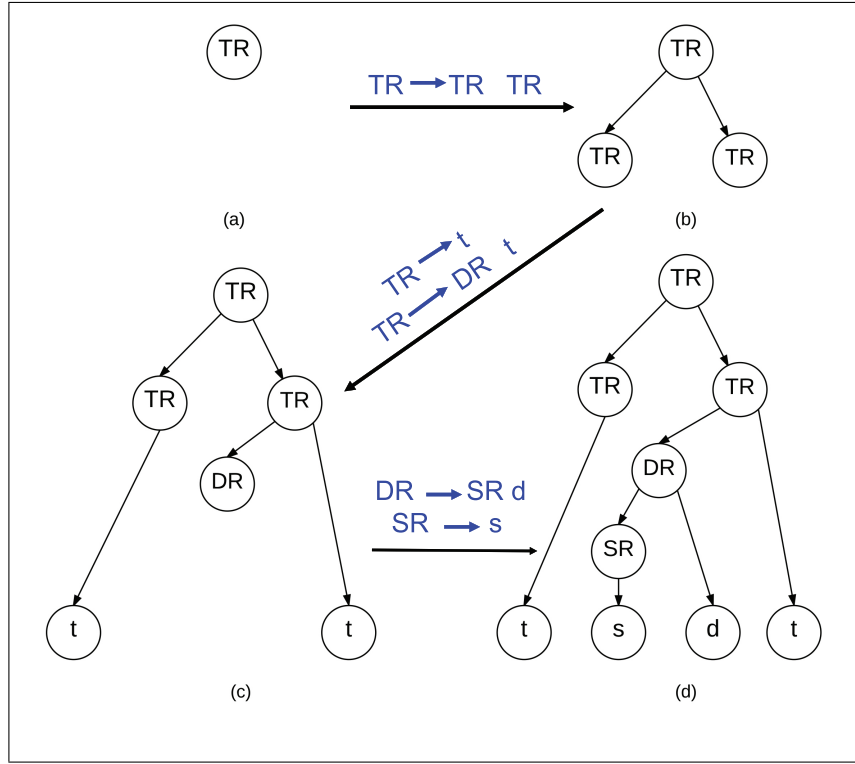
*Voice leading* captures the melodic attraction between consecutive chords per voice. Voice leading is a horizontal measure that estimates the fitness of each note in a voice according to the previous note. Theoretically, the number of semitones, the stability of the notes, and the overall fitness between the chords can affect the evaluation of the voices [9, 14]. Lerdahl [9] states that the stability of each note is related with the distance to the tonal center or key. The closer to the key, the more stable the note is.

*Hierarchical structure* captures the multiple levels of the musical structure by adopting tree-based structures driven from functional harmonic categories. Lerdahl’s hierarchical approach analyzes the hierarchical structure through the tonal tension measure, which can be divided into local tension and inherent tension to know how the chords are hierarchically related. Local tension measures the distance between the chord we want to evaluate and the chord immediately above it in the tree, called parent chord. Likewise, inherent tension considers all the distances between the parent chord and all the chords above.

### 2.3 Rohrmeier’s Hierarchical Tree Structure

Rohrmeier’s computational approach to encode the hierarchy of chord progression is detailed, towards the definition of an evaluation scheme for hierarchical tonal fitness. Rohrmeier applies principles from theories focusing on the hierarchical dimension of music using a binary grammar that respects tonal rules to generate chord progressions and explicitly designed to be computationally feasible and testable [16]. His hierarchical model of tonal music relies on four categories, of which the functional level is of relevance here. Consequently, we apply this generative grammar to our work to measure the hierarchical structure.

According to Rohrmeier, the rules in eq. (4a) to eq. (4g) characterize the core behavior of a grammar which represents the functional regions in a chord progression. The rules contain two kinds of symbols: *non-terminal symbols* represented by capital letters and *terminal symbols* represented by lower-case letters. The *non-terminal symbols* represent functional regions such as the tonic region TR, the dominant region DR, the subdominant region SR, and any of the previous regions XR. Non-terminal symbols can be expanded into different harmonic regions following the rules in eq. (4a) to eq. (4d) or into terminal symbols following eq. (4e) to eq. (4g). The *terminal symbols*  $t$ ,  $d$ ,  $s$  represent tonic, dominant, or subdominant chords in the sequence respectively, and cannot be further ex-



**Fig. 1.** Representation of the construction of a hierarchical structure for a three-chord progression.

panded.

$$TR \rightarrow TR \ DR \tag{4a}$$

$$TR \rightarrow DR \ t \tag{4b}$$

$$DR \rightarrow SR \ d \tag{4c}$$

$$XR \rightarrow XR; \quad XR \ \forall \text{ non-terminal} \tag{4d}$$

$$TR \rightarrow t \tag{4e}$$

$$DR \rightarrow d \tag{4f}$$

$$SR \rightarrow s \tag{4g}$$

Figure 1 shows an example of how to construct a tree structure. The first node in Figure 1(a) is a tonic region  $TR$ , which is responsible for defining the key of the chord progression. Firstly, we apply Rule (4d) for the tonic region. Then, the first child  $TR$  applies the Rule (4e) and the second child applies (4b) (Figure 1(c)). The final Figure 1(d) is obtained by applying Rules (4c) and (4g).

### 3 Measuring the Tonal Fitness of a Chord Progression

The properties of the TIS derived from mathematical measures such as the Euclidean distance or the angle between TIVs are the main indicators used to measure the tonal fitness of a pitch configuration [3]. In the present work, we aim to create a measure that captures the tonal fitness of a chord through its musical properties. Following Lerdahl’s proposal, we create a measure in eq. (5), where each term corresponds to the codification of an element proposed in Sec. 2 but encoded in the TIS: *tonal pitch distance* ( $\delta$ ), *surface dissonance* ( $c$ ), *voice leading* ( $m$ ), and *hierarchical structure* ( $h$ ):

$$M(T_i, P) = k_1\delta(T_i, P) + k_2c(T_i) + k_3m(T_i, P) + k_4h(T_i, P), \quad (5)$$

where  $T_i$  is the  $i^{th}$ -chord of the progression  $P$  and  $k_j$  are constants that represent the weights of each parameter. The following subsections will detail how the four items are encoded in the TIS so that we can measure them mathematically.

#### 3.1 Tonal Pitch Distances

The measure  $\delta$  encodes the tonal pitch distance between two consecutive chords. According to Section 3.1,  $\delta$  measures three musical properties: distance to the previous chord in the sequence, distance to the key, and distance to the tonal function. In eq. (6),  $\delta$  is decomposed as

$$\delta(T, P) = \rho(T_i, T_{i-1}) + \theta(T_i, T_{key}) + \theta(T_i - T_{key}, T_f), \quad (6)$$

where  $\rho(T_i, T_{i-1})$  uses eq. (1) and  $\theta$  uses eq. (3).  $\theta(T_i, T_{key})$  measures the degree of membership of  $T_i$  to the main key of the full chord progression and  $\theta(T_i - T_{key}, T_f)$  measures the similarity to the tonal function given by the tree previously built.

In the TIS, chords musically close to the key have small distances between the key configuration and the chord configurations. The key is represented by  $T_{key}$  obtained as the TIV of the chroma vector  $C_{key}(n)$  that contains all the notes of the scale corresponding to the key. Then, eq.(3) estimates the angle between the chord  $T_i$  and the center of the key  $T_{key}$ .

To measure the proximity of one chord to a tonal function in the TIS, we check if the chord  $T_i$  is aligned with the tonic I, subdominant IV, or dominant V by using the angle measure proposed in eq. (3).  $T_i - T_{key}$  is the chord using  $T_{key}$  as reference and  $T_f$  is a vector representing one of the harmonic functions I, IV, and V also referenced by  $T_{key}$ . We aim to minimize  $\theta(T_i, T_{key}, T_f)$  using  $T_f$  following the harmonic sequence given by the tree built in the previous step to ensure  $T_i$  is aligned with one of them.

#### 3.2 Consonance

In the TIS [3], the norm  $\|T_i\|$  measures the consonance  $c$  of the chord represented by  $T_i$ . If  $c$  results in large values of  $\|T_i\|$ , the chord is very consonant. Therefore, we aim to maximize  $c(T_i)$ .



### 3.3 Voice Leading

The measure  $m$  represents the melodic attraction of two consecutive chords. Lerdahl proposes three factors to model the melodic attraction, which are the perceptual distance between the chords, the number of semitones between the notes and the stability of the notes (how much they attract the rest of the notes). Following this, the voice leading measure between  $T_i$  and  $T_{i-1}$  has been encoded as

$$m(T_i, P) = \frac{\sum_{l=1}^3 v(n_{l_{i-1}}, n_{l_i})}{\rho(T_i, T_{i-1})}, \quad (7)$$

where  $\rho$  is the Euclidean distance from eq. (1) which captures the perceptual distance between the chords and  $v$  is a measure of voice-leading for each voice  $l$  produced between the note  $n_{l_i}$  of the present chord and the note  $n_{l_{i-1}}$  of the previous chord. To calculate the stability of the notes of a chord in a progression, we measure the distance of its corresponding pitch class to the key in the TIS. The number of semitones and the distance to the key are calculated as

$$v(n_{l_i}, n_{l_{i-1}}) = \rho(T_{n_{l_i}}, T_{key})e^{0.05s} \quad (8)$$

where  $s$  is the number of semitones between  $n_{l_i}$  and  $n_{l_{i-1}}$ .

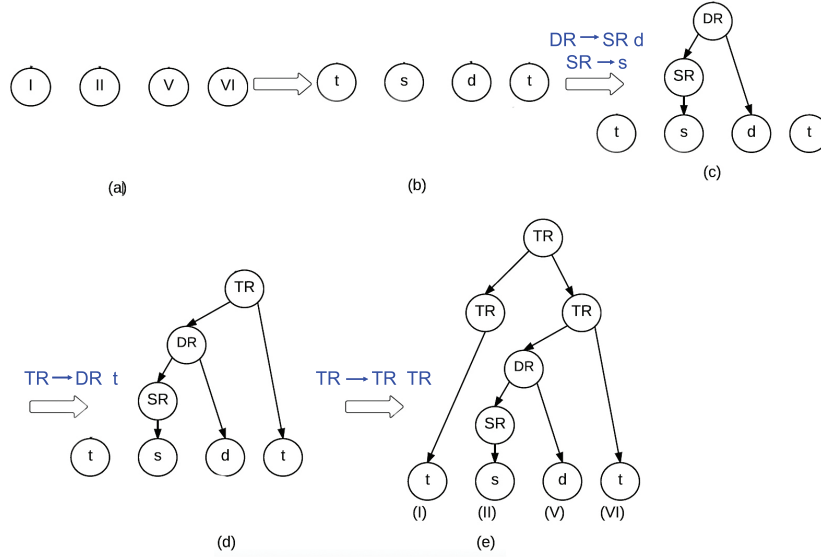
### 3.4 Hierarchical Tension

The value of  $h$  represents the tonal tension concept following Lerdahl's theory, which considers distances between chords that are closer in a tree structure modeled for a progression. The tension related to the hierarchical structure of the progression is calculated as

$$h(T_i, P) = \rho(T_i, T_j) + \frac{\sum_{k=j}^N \rho(T_i, T_k)}{N}, \quad (9)$$

where  $N$  is the number of chords in the tree and  $T_j$  is the parent chord of  $T_i$  and  $T_k$  is the parent chord of  $T_{k-1}$  following the tree structure of the chord progression  $P$ . Here,  $\rho$  is the Euclidean distance between two chord codifications calculated with eq. (1).

Rohrmeier *et al* [16] proposes a generative grammar to create hierarchical structures following tonal music principles. However, Rohrmeier's work is not oriented to generate chord progressions computationally based on his grammar. Given a progression, we need to extract one tree that represents the hierarchy between the chords in the progression. Therefore, it is necessary to implement a method that calculates the node in the tree corresponding to the specific chord considered by inversely applying the rules of Rohrmeier's proposal [16]. Firstly, the leaf chords are always replaced by  $t$ ,  $s$  or  $d$ . To decide which harmonic function is aligned with each chord, we calculate the angle between the tonal function and the chord in the TIS. In a second step, we apply the grammar rules inversely following a hierarchy according with three criteria:

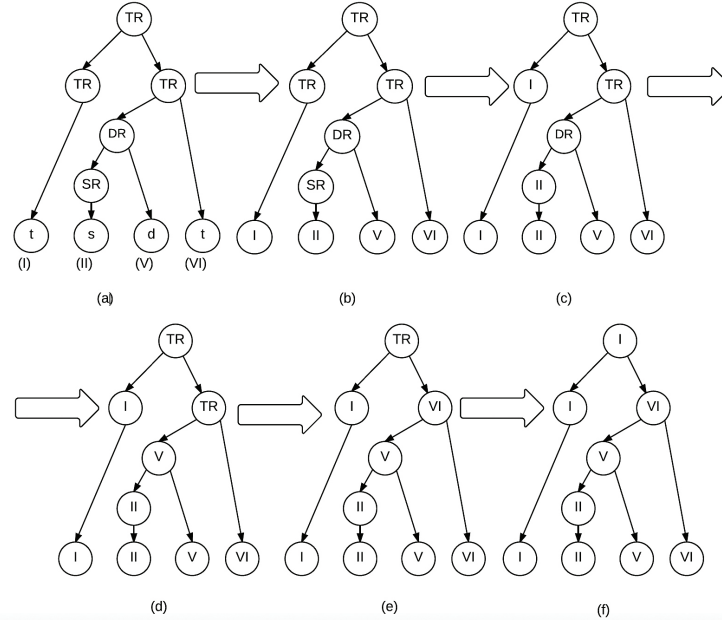


**Fig. 2.** Process to obtain a tree structure from a chord progression.

- To avoid unfeasible trees, select first the element ‘s’, which appears in a lower number of rules.
- To avoid unfeasible trees, try to build the tree from the inner leaves and to connect them with the outer ones.
- To avoid trees with too many nodes, always try to apply a rule which implies a greater number of elements, if possible. In case we have several rules that accomplishes this criterion, select one randomly.

To clarify this process, a simple example considering the tree of Figure 1 is illustrated in Figure 2. In the first stage, the chords are replaced by ‘t’, ‘s’ or ‘d’, according to their alignment with the harmonic functions. Secondly, as we have an ‘s’ element and this can be connected to the ‘d’ element, so we apply Rule (4c) (Figure 2(c)). Now, we can connect the first ‘t’ element, or the last ‘t’ element. Randomly, we selected the final ‘t’ and applied inversely Rule (4b). Finally, we connected the first ‘t’ with Rule (4e) and Rule (4d) (Figure 2(d)).

With the hierarchical tree, the tension can be now calculated. However, we still need to know the importance of a chord in the hierarchical structure and consequently their local tension and inherent tension. To calculate their tension, we need to know which chords dominate (have more tension) the other chords.



**Fig. 3.** Visualization of a tree structure with specific chords to measure the tension.

The hierarchy of “dominated” chords is represented by replacing each node of the tree constructed previously with the most tense chord, selected from its children. We implemented a method to obtain a tree with the “dominant” chords of each level, and therefore, to be able to calculate the local and inherent tensions. Firstly, the leaf nodes are replaced by the specific chords. The parent nodes with only one child automatically represent the leaf chord. The parents with children representing different tonal regions are automatically represented by the child with the same harmonic function that the parent. Finally, the parent with children of the same tonal regions is replaced by the most stable chord according to the measure described in Section 3.1. A simple example considering the tree of Figure 1 is illustrated in Figure 3. Step a represents the initial tree and the chord progression. Step b replaces the terminal by the degrees. Step c replaces the *SR* and *TR* nodes with the only child they have. Step d replaces the *DR* with *V*. Similarly, step e replaces *TR* with the tonic chord *V* and step f chooses the most stable chord according to the tonal distance proposed in the previous section, selecting *I* as the parent node.

## 4 Evaluation

The evaluation aims to demonstrate that  $M$  reflects the *tonal fitness* of a chord in a progression. Firstly, we need to evaluate  $M(T_i, P)$  in optimal conditions. That means we need to find the best values for weights  $k_j$  in eq. (5) by applying

cross validation. In a second step, we validate  $M(T_i, P)$  by comparing subjective ratings collected by a listening test and the objective measure for different chord progressions.

Two premises were considered to create the listening test. Firstly, the measure  $M$  should be able to capture tonal fitness for both *fit* and *unfit* chords. Thus chords with a high level of *tonal fitness* would have low values of  $M$  and vice-versa. Secondly,  $M$  should reflect their tonal fitness independently of the main key and of the hierarchical tree. Additionally, we have to consider that the evaluation of harmony could be influenced by multiple factors. To avoid external elements that can bias the subjective, the listening test contains controlled chord progressions in a basic musical context with specific and short chord progressions.

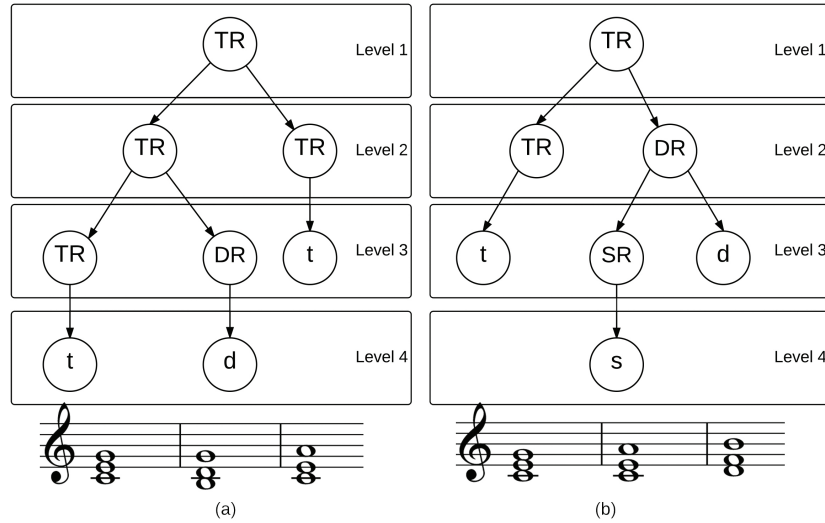
The listening test presented a sequence of three three-note chords and asked the listener to rate how well the last chord fits the progression. Note that the three-note chords can be triads or contain non-harmonic tones. To demonstrate that  $M(T_i, P)$  reflects the *tonal fitness* of chords, we selected triad chords randomly but with objective ratings that sample the function  $M$  from low values (chords with a high *tonal fitness*) to high values. The first chord in each progression was always the tonic in root position to establish the key [1] because the tonic determines the tonal basis of the music. Additionally, the root position triads have a firmer sense of tonal centering, resulting in the difference in pitch between the major and minor modes. The second chord is always different for each key to represent different harmonic functions. Additionally, to demonstrate its independence of tree structures and keys, the chord progressions were classified in four groups, two for a tree with the sequence Tonic-Dominant-Tonic, in G major and C minor, and two for a tree with the sequence Tonic-Subdominant-Dominant, in C minor and G minor. Both tree structures with some examples of chord progressions for each tree are shown in Figure 4.

The listening test is online<sup>6</sup> and consists of four playlists with the chord progressions presented randomly. Each chord progression can be played multiple times before assessing it. The listeners were asked to evaluate how well the third chord follows the first and second using the following ratings: very good (+2), good (+1), fair (0), bad (-1), or very bad (-2). In total, 48 people took the test, among which ten declared no musical training, nineteen considered themselves amateurs, and twenty were professional musicians.

We expect the objective measure  $M(T_i, P)$  to correlate well with the ratings from the subjective evaluation and reflect the *tonal fitness* of each chord because  $M(T_i, P)$  includes information from the linear and hierarchical dimensions. The sum of the distances in the hierarchical tension can change according to the tree design, so we expect that a tree can influence the correlation between the objective measure and the subjective ratings.

---

<sup>6</sup> <http://form.jotforme.com/form/52522142163343>



**Fig. 4.** Representation of the construction of a hierarchical structure for a three-chord progression.

## 5 Results and Discussion

$M$  is a weighed linear combination of four terms with weights  $k_1$ ,  $k_2$ ,  $k_3$  and  $k_4$  that determine the influence of each corresponding term. We used cross validation to calculate the weights  $k_i$  that best fit the scores resulting from the listening test. Among the total of 30 chords for each tree structure that were rated by the listeners, 24 chords were selected to training the 5-fold cross validation, while the rest (6 chords) were applied for the validation part.

For a particular fold, the set of training sequences might well contain sequences very like the test sequences. To minimize this possibility and increase the reliability of the validation, we calculated the weights and the statistical measures for all the possible combinations of 30 chord progressions of the same tree structure, which makes a total of 4060 possible combinations. The final weights obtained were  $k_1 = 4.22$ ,  $k_2 = 2.13$ ,  $k_3 = 2.06$ ,  $k_4 = 3.76$ .

Once the weights are incorporated into the measure  $M$  for the experiment, we try to validate if the measure  $M$  captures the *tonal fitness* of a chord in a progression. In other words, we are evaluating if lower values of  $M$  are associated with chords that scored higher in the listening test. Table 1 shows the statistical results (linear regression and error) of the subjective ratings versus the measure  $M$  for the chords included in the listening test.

The first column of the table shows the linear regression of the subjective ratings versus the measure  $M$ ,  $R^2$ , along with statistical values calculated from the data, ( $p$ -value). The first row contains the results for the chord progressions with the hierarchical structure shown in Figure 4(a). The second row presents the

results for the chord progressions with the tree structure shown in Figure 4(b). The statistical analysis shows that  $M$  captures the tonal fitness of the chords. The  $p$ -values are all below the 1% threshold for the null hypothesis, the high  $R^2$  values indicate that the subjective ratings correlate strongly with the objective values.

Likewise, we also include a comparison between our measure  $M$  and  $L(T_i, P)$  proposed by Lerdahl [8] and  $D(T_i, P)$  from a previous work [11]. Both  $L(T_i, P)$  and  $D(T_i, P)$  are designed as a linear combination of different terms that encode musical factors.  $L(T_i, P)$  in the second column considers *consonance*, *tonal tension*, *melodic attraction* or *voice leading* and *distance to previous chord* to calculate the *tonal fitness* of a chord in a different representation space. Thus, the goal of the comparison with  $L(T_i, P)$  and  $M(T_i, P)$  is to investigate how the representation space can affect to the correlation values between subjective ratings and the evaluation measure. In turn,  $D(T_i, P)$  in the third column considers only the *consonance*, the *distance to previous chord* in the progression, the *distance to the key* and the *distance to the tonal function* [11] in the TIS.  $D(T_i, P)$  measures the *tonal fitness* of a chord  $T_i$  by considering the linear dimension, ignoring hierarchical relationships between the chords. Therefore, we aim to analyze how the hierarchical and linear dimensions encoded in  $M$  can influence the subjective ratings against considering only the linear dimension encoded in  $D$ , in the same representation space.

We used the results of the same listening test to calculate the optimal weights of each term of  $L$  and  $D$  via cross validation. The statistical results of  $L$  and  $D$  are shown in columns two and three of Table 1, respectively, and are also grouped by the tree structure of the chord progressions considered.

The statistical analysis shows that  $M(T_i, P)$  captures the *tonal fitness* of the chords better than  $L$  and  $D$ . The high  $R^2$  values obtained indicate the linear regression of the subjective ratings as a function of the objective values and the  $p$ -values are all below the 1% threshold for the null hypothesis, which suggests that the listeners' rates fit the objective measure well. Note that Lerdahl's measure  $L(T_i, P)$  also correlates with the subjective ratings of the listening test, but obtaining slightly lower values for  $R^2$  and  $p$ -value. Likewise, the  $R^2$  and  $p$ -value in  $D(T_i, P)$  are correlated, but below the rates obtained for  $M$ .

A limitation of the experimental setup is that the listening test was designed to evaluate the chord in a short harmonic context with only two previous chords. Consequently, a three-chord sequence may be perceived by some listeners as a

**Table 1.** Comparative table of the statistics  $R^2$  and the  $p$ -value for the measures  $M(T_i, P)$ , Lerdahl's  $L(T_i, P)$  and [11]  $D(T_i, P)$

		$M(T_i, P)$	$L(T_i, P)$	$D(T_i, P)$
T.1	$R^2$	$0.94 \pm 0.10$	$0.69 \pm 0.45$	$0.90 \pm 0.21$
	$p$ -value	$0.03 \pm 0.15$	$0.39 \pm 0.23$	$0.17 \pm 0.18$
T.2	$R^2$	$0.94 \pm 0.08$	$0.59 \pm 0.43$	$0.85 \pm 0.27$
	$p$ -value	$0.07 \pm 0.13$	$0.41 \pm 0.26$	$0.17 \pm 0.16$

cadential progression, and the listener may expect chords that somehow solve the tension of the previous chords. A longer progression could be needed to reach stronger conclusions about the *tonal fitness* and how the harmonic context and the hierarchical structure can influence the musical perception of one particular chord.

Another limitation is the tree construction of this approach, which does not consider rhythmic features. Rhythm sometimes can influence the hierarchical perception of which chord is above others in the structure, and therefore can also influence the musical perception of the listeners. Taking these factors out of the equation means that some trees could be theoretically equivalent, but in practice, musically perceived in different ways, and consequently, subjectively rated differently. We plan to investigate how the construction of different trees for a particular chord progression can influence the ratings in a future work.

Despite the limitations of this experiment, one of the contributions of this measure is the beginning of a generalization function that can study a long-term chord progression based on both hierarchical and linear structures in different tonalities. The adaptation of Lerdahl’s space to the TIS to explain the tonal properties of a chord besides tonal tension and voice leading can be also considered a contribution for the community. The positive results are encouraging to apply this measure to real analysis problems and computational generative systems.

## 6 Conclusions and Future Work

In this work, we designed a model that captures the *tonal fitness* of one chord in a particular chord progression encoded in the Tonal Interval Space [3]. We use the work proposed by Rohrmeier [16] to represent the hierarchical structure associated with a chord progression. Then we measure the *tonal fitness* of each chord considering factors such as consonance, the distance from the previous chord, the distance to the key, voice leading, and the hierarchical tension of the chord with respect to the rest of the progression. We performed a listening test to evaluate how well chords fit a progression for two different tree structures in different keys. Statistical analysis showed that the subjective ratings correlate strongly with the objective values, validating the objective measure as a proxy for the subjective measure of *tonal fitness*. We also compared the present model with Lerdahl’s [8] and a previous proposal [11]. In this preliminary experiment, the measure proposed here was a good indicator of tonal fitness independently of the tree resulting from the hierarchical structure.

Future work will analyze how longer progressions can influence the results obtained with a new listening test, and how different tree structures affect the subjective results. Likewise, the measure can be applied not only to analyze chord progressions from real music samples, but also can be included in a generative system of chord progressions. In a future work, we will investigate how to improve a system to generate chord progressions using an artificial immune system described in a previous work [11] with this new measure to create chord

progressions that follow a given pattern, like a tension curve or a given structural tree.

## 7 Acknowledgments

Research partially funded by the project “Co-POEM:Platform for the Collaborative Generation of European Popular Music” (ES01-KA201-064933) under the program Erasmus+:KA201-Strategic Partnership; by the project “Experimentation in music in Portuguese culture: History, contexts and practices in the 20th and 21st centuries” (POCI-01-0145-FEDER-031380) co-funded by the European Union through the Operational Program Competitiveness and Internationalization, in its ERDF component, and by national funds, through the Portuguese Foundation for Science and Technology; and by the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 831852 (MORPH).

## References

1. Agmon, N.: The grotthuss mechanism. *Chemical Physics Letters* **244**(5), 456–462 (1995)
2. Amiot, E.: *Music Through Fourier Space*. Springer (2016)
3. Bernardes, G., Cocharro, D., Caetano, M., Guedes, C., Davies, M.E.: A multi-level tonal interval space for modelling pitch relatedness and musical consonance. *Journal of New Music Research* **45**(4), 281–294 (2016)
4. Bernardes, G., Cocharro, D., Guedes, C., Davies, M.E.: Harmony generation driven by a perceptually motivated tonal interval space. *Computers in Entertainment (CIE)* **14**(2), 6 (2016)
5. Bernstein, L.: *The unanswered question: Six talks at Harvard*. Harvard University Press (1976)
6. Callender, C., Quinn, I., Tymoczko, D.: *Generalized chord spaces*. Princeton University, Department of Music (2006)
7. Herremans, D., Sørensen, K., Martens, D.: Classification and generation of composer-specific music using global feature models and variable neighborhood search. *Computer Music Journal* **39**(3), 71–91 (2015)
8. Lerdahl, F.: *Tonal pitch space*. Oxford University Press (2005)
9. Lerdahl, F., Krumhansl, C.L.: Modeling tonal tension. *Music Perception: An Interdisciplinary Journal* **24**(4), 329–366 (2007)
10. Müllensiefen, D., Wiggins, G.: Polynomial functions as a representation of melodic phrase contour. *na* (2011)
11. Navarro-Cáceres, M., Caetano, M., Bernardes, G., de Castro, L.N.: Chordais: An assistive system for the generation of chord progressions with an artificial immune system. *Swarm and Evolutionary Computation* (2019)
12. Pachet, F.: The continuator: Musical interaction with style. *Journal of New Music Research* **32**(3), 333–341 (2003)
13. Quick, D., Hudak, P.: *A temporal generative graph grammar for harmonic and metrical structure*. Ann Arbor, MI: Michigan Publishing, University of Michigan Library (2013)



14. Quinn, S., Watt, R.: The perception of tempo in music. *Perception* **35**(2), 267–280 (2006)
15. Riemann, H., Hagh, R.H.: *History of music theory*. University of Nebraska (1962)
16. Rohrmeier, M.: Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music* **5**(1), 35–53 (2011)
17. Schenker, H., Oster, E.: *Free Composition: Volume III of New Musical Theories and Fantasies*. Pendragon Press, Hillsdale (1979)
18. Schoenberg, A.: *Theory of harmony*. Univ of California Press (1978)
19. Steedman, M.: The blues and the abstract truth: Music and mental models. *Mental models in cognitive science* pp. 305–318 (1996)
20. Whorley, R.P., Wiggins, G.A., Rhodes, C., Pearce, M.T.: Multiple viewpoint systems: Time complexity and the construction of domains for complex musical viewpoints in the harmonization problem. *Journal of New Music Research* **42**(3), 237–266 (2013)
21. Woolhouse, M.: Modelling tonal attraction between adjacent musical elements. *Journal of New Music Research* **38**(4), 357–379 (2009)