# AuToTen: a video introduction - Transcript

## Intro slide

Hi everyone, my name is German Ruiz-Marcos, and today I’m presenting a research project I have been working on together with Alistair Willis and Robin Laney at the Open University. As you can see, this project concerns the automatic calculation of tonal tension.

## Slide 1

In the literature, many authors have identified the concept of musical tension as essential to the listening experience. As a consequence many scholars have attempted to model musical tension. These models normally study the relationship or correlation between different features and musical tension. These features can be either psychoacoustic, such as loudness, tempo changes or onset frequency, among others; or can be cognitive features, which mainly concern melodic and harmonic motion. Models based on psychoacoustic features are, in general, easy to automate. You can probably think of different existing tools which automatically calculate loudness levels or the onset frequency of a given audio signal. On the contrary, models based on cognitive features are, most of the time, theoretical. We believe this lack of automated cognitive models of tension is a gap that needs to be filled, and so we have been working on this project.

## Slide 2

Taking into account that we are focussing on cognitive models of tension, let me borrow Lerdahl and Krumhansl’s definition of tonal tension, which they describe as the sense of tension created by melodic and harmonic motion. From the existing models of tonal tension, we believe Lerdahl’s is the most complete, and that is because of five reasons.

## Slide 3

First, it considers rhythmic, melodic and harmonic components, as opposed to other existing models which just focus on one or two of these components at a time. Second, it takes into account a piece’s hierarchical structure, as opposed to other existing models which only calculate tension following the chords in a piece of music in a sequential way. Third, it produces quantitative values of tension which are valuable for researchers in music computing like us, as we can use these values to implement new computational tools. Fourth, its components connect with other cognitive theories, such as those of Leonard Meyer, E. Narmour, Steve Larson, J. Bharucha or C. L. Krumhansl, among others. And fifth, it has shown strong correlations against human judgements of tension.

## Slide 4

Let me briefly introduce how Lerdahl’s model of tonal tension, or MTT, works. First, we need to select a piece of music. Then, we need to calculate its key and chord labels. Next, we need to calculate the piece’s hierarchical structure. To do so, we need to use Lerdahl and Jackendoff’s Generative Theory of Tonal Music, or GTTM. This theory consists of a collection of rules to extract the metrical components of a piece of music, its inner groups and patterns, and its hierarchical relations. Finally, we need to apply the rules in Lerdahl’s MTT. MTT incorporates two different components of tonal tension. The first, harmonic tension, which in the literature is referred to as Tglob, concerns the vertical arrangement of chords and the cognitive distances between them. The second, attraction, which in the literature is referred to as alpha (α), concerns the horizontal arrangement of chords and the voice leading paths between chords. To calculate the values of harmonic tension of a given piece of music, we need to calculate the cognitive distances between the chords connected by the piece’s hierarchical structure. To do so, MTT relies on Lerdahl’s Tonal Pitch Space, or TPS, which includes a quantitative method to calculate the distance between chords in a given key, that might be perceived by Western listeners. Likewise, we need to calculate a collection of surface parameters, which concern three features: the chord’s inversions; the chord’s highest notes; and the role the notes in each chord play in the context of the piece’s key. To calculate the values of attraction, we need to parse each of the inner melodies included in every chord transition, as well as account for the intervals within these melodies. And, we need to use TPS’s distance calculation again. Let me show you an example of the application of Lerdahl’s model of tonal tension.

## Slide 5

So, MTT’s first step is to pick a piece of music. Here we have the Grail theme from Wagner’s Parsifal. I have chosen this piece because you can find a detailed explanation of what I’m about to say in the literature. The second step concerns the piece’s harmonic analysis. As you can see at the bottom line, the theme is in the key of e-flat major, and its chord’s labels consist of a very common diatonic harmonic progression. Next, we move into the piece’s hierarchical structure. By applying GTTM’s rules on the Grail theme, one will get the tree shown on top of the picture. This tree illustrates the importance of each note with regards to the whole theme. Finally, by applying MTT’s rules, one can calculate the quantitative values of harmonic tension and attraction, which are shown in the middle space. The question now is, well, is there a way of calculating these values automatically?

## Slide 6

To answer the question, let me introduce AuToTen, as in Automatic Tonal Tension, which is in fact, the core of my talk. AuToTen is a system we have developed to automatically calculate the contributions to tonal tension of a piece of music according to Lerdahl’s Model of Tonal Tension. As input, AuToTen needs to be fed with a piece of music as well as with the outputs of its GTTM analysis. All these files are in MusicXML format. In our project, all GTTM analysis files were calculated using the interactive GTTM analyser, developed by Hamanaka and colleagues, and which is publicly available. See the reference in the footnote for more detail. As output, AuToTen will produce the quantitative values of harmonic tension and attraction, according to Lerdahl’s Model of Tonal Tension, like those we just saw in the previous slide.

## Slide 7

So, how does AuToTen work? For the sake of simplicity, I’m going to indicate how the steps one needs to take to apply Lerdahl’s Model of Tonal Tension, those which I have already described, are implemented in AuToTen. First, the piece of music to be analysed needs to be input by the user. Second, to calculate the piece’s key and chord labels, AuToTen uses the toolkit Music21, being developed by MIT. More specifically, it uses the modules Analyse and Chordify. The hierarchical structure is another input file. However, AuToTen transforms the input GTTM tree into a matrix to ease the upcoming calculations. Recall that these GTTM related files can be calculated by the user using the interactive GTTM analyser. To calculate the piece’s harmonic tension, we have transformed Lerdahl’s Tonal Pitch Space components into different objects and its rules into functions which combine different classes. In this way, AuToTen is capable of calculating the cognitive distance between chords. Likewise, the surface parameters (those concerning the chords’ inversions and their melodic notes) are calculated by extracting data on the pitches using Music21 and mapping these data with the weightings defined in Lerdahl’s MTT. Finally, to calculate the piece’s values of attraction, we have represented the chords’ inner melodies numerically, to account for every transition and their respective influence.

## Slide 8

To evaluate AuToTen, we asked ourselves the following questions. Has Lerdahl’s MTT been correctly implemented, and how accurate are the calculations of tension provided by AuToTen? Although these questions may seem pretty similar they are looking at very different things. With the first question, we wanted to check whether we had made any mistakes when implementing Lerdahl’s Model of Tonal Tension computationally. We carried out 100 test cases to give an affirmative answer to the first question. However, as I don’t have much time, I am not presenting these test cases in today’s talk, so please refer to the paper and to AuToTen’s repository. With the second question we wanted to check, how accurate were AuToTen’s quantitative outputs compared to those calculated when manually using MTT. You may be wondering why AuToTen might not be 100% accurate, now that we have seen after answering the first question that it has been correctly implemented. However, it should be noted that AuToTen relies on the accuracy of the GTTM trees provided as input (recall that we took these from the interactive GTTM analyser), as well as on the accuracy of the chord labels which were calculated using Music21. As you can imagine, neither present an accuracy of 100%, and so we ask ourselves the second question.

## Slide 9

To answer this question, we have carried out three tests using four pieces of music. But, why just four pieces of music, and why these four? As it has already been stressed in the literature. There exists a lack of tension-related databases. In the case of Lerdahl’s Model of Tonal Tension, these four pieces are the ones that have been analysed and studied by Lerdahl and Krumhansl in the past, thereby these are the most reliable and accessible data one could find when working on Lerdahl’s Model of Tonal Tension. Using these data, we have carried out three accuracy tests. A first test to study the accuracy of the chord labels found by AuToTen compared to those in the groundtruth. A second test to study the accuracy of the GTTM trees used as input in AuToTen compared to those in the groundtruth. And a third test to study the correlation between AuToTen’s outputs and those given by the groundtruth. Again, because I have not much time, I will be only focussing on the third test. You can find more detail about the other two tests, their statistics and more data on the paper and in AuToTen’s repository.

## Slide 10

Before we get into AuToTen’s evaluation, we need to discuss how Lerdahl’s MTT was evaluated. Lerdahl and Krumhansl recruited a group of participants and asked them to judge the degree of tension they perceived when listening to the pieces of music shown in the previous slide. As an example, this figure includes a dotted line with the average values of tension perceived by the participants when listening to the Grail theme from Wagner’s Parsifal, which recall has already been introduced in this presentation. Using these data as groundtruth, Lerdahl and Krumhansl calculated a multiple regression to find the most accurate prediction of tonal tension that MTT could give. See how the predictive values of tonal tension were seen as the dependent variable, and MTT’s harmonic tension and attraction as independent variables. The results include an *R2*value as a goodness of fit measure, between in parentheses you can find the degrees of freedom, first the number of independent variables in the regression, and second the number of events in the piece of music minus two. They also calculated an adjusted value of *R2*to make it more comparable with other models for the same data that have different numbers of degrees of freedom. Likewise, they calculated the regression coefficients (the betas), and the statistical significance given by the *p* values. In this example concerning the Grail theme, we can see MTT’s prediction of tonal tension (the blue line in the figure) strongly correlate with the degrees of tension perceived by human listeners and are all statistically significant.

## Slide 11

To test the accuracy of AuToTen’s outputs, we have followed a similar methodology. However, in this case the groundtruth data are the predictions calculated by MTT in the previous slide, which are shown again with a blue line in this new figure. The independent variables correspond to the measures of harmonic tension and attraction calculated by AuToTen. In this way, the predictions of AuToTen that best fit MTT’s groundtruth are those shown by a dotted line in the figure. I will be showing the data corresponding to this multiple regression in the next slide.

## Slide 12

This slide shows the results of the previous multiple regression, the one concerning Wagner’s theme, as well as the results of the rest of the pieces for similar calculations. For each piece of music, there is the value of *R2* and the degrees of freedom, *n* and *m*, an adjusted value of *R2* and the *p* values corresponding to harmonic tension and attraction calculated by AuToTen. By looking at the *R2*values, that is the first column, it can be seen our regression shows strong correlation but that of Chopin’s, whose correlation is healthy, and are all statistically significant as indicated by all *p* values included in the last two columns. Unlike the rest of the pieces, Chopin’s is constantly modulating through different keys. In fact, the prelude’s last two phrases transition across seven keys. We believe these heavy modulation processes hindered the automatic calculation of chord labels performed by Music21, which in turn affect the final valuations of tension and could explain the lower accuracy in the case of Chopin’s piece. Likewise, notice how, as the number of events increases (see Bach’s and Chopin’s rows), the correlation tends to decrease because there are more possible points of deviation, as indicated by *m* (the second degree of freedom). In addition to the regressions shown in this slide, we also calculated other regressions to study the correlation of harmonic tension and attraction independently. Please check the paper if you are interested in these data and the respective discussions.

## Slide 13

To sum up, what is the message you should be taking home from this talk? Well, simple and easy. AuToTen has been launched. It incorporates an automation of Lerdahl’s theories, it is publicly available, and it has shown strong correlations against groundtruth data. In future, we intend to produce a more detailed report of AuToTen’s implementation. We believe the community will benefit from a thorough discussion about the computational challenges encountered when automating Lerdahl’s theories. Likewise, we have already started building a dataset of tonal tension based on AuToTen, which we believe can have a great impact in the community. Finally, we aim to explore the automatic generation of music with regards to tension using the doors opened by AuToTen, such as accounting for a piece’s hierarchical structure.

## End slide

And that’s it, many thanks.