

# Singing from the same sheet: A new approach to measuring tune similarity and its legal implications

Daniel Müllensiefen  
Department of Psychology  
Goldsmiths  
University of London

Robert J.S. Cason  
School of Law  
Birkbeck  
University of London

# Outline

1. Introduction
2. Legal aspects of the Project
3. Similarity in Copyright Disputes
4. The Legal Degree of Similarity
5. The Substantial Part Doctrine
6. Studying Music Similarity Empirically
7. Two Perceptual studies and a Computational Model
8. Legal Implications of the Computational Model

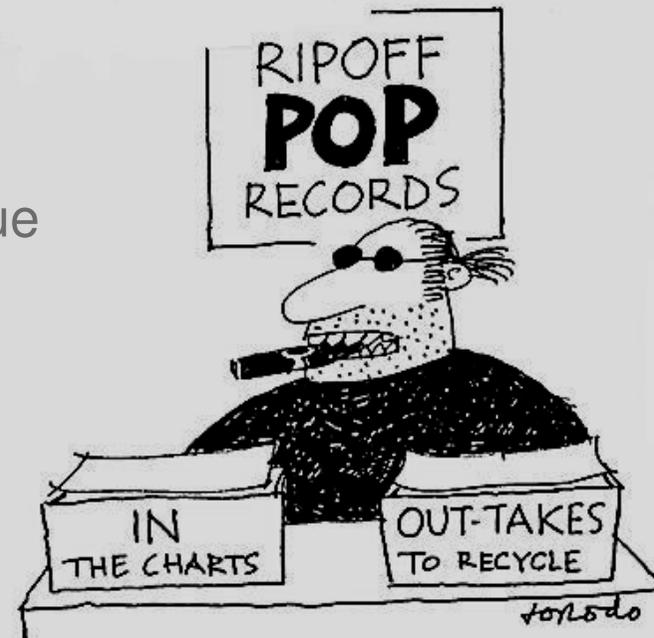
# The Legal Side of the Project

- To explore if court decisions on music plagiarism could be accurately predicted by formal models
- Case law from commonwealth countries
- First comprehensive case law database of melodic infringement disputes in commonwealth countries
- Database is currently hosted at the Music Copyright Infringement Resource UCLA



# Introduction

- Music Plagiarism
  - High Commercial Interest
  - Captivates the interest of the public
  - Simplistic/Repetitive nature of pop music
- Little research into the potential use of musical comparison technologies for copyright disputes
- Represents a new interdisciplinary angle in which to analysis and critique the law



# Similarity in Copyright

- Altered or ‘non-identical’ copying of a part
- Copyright Design and Patents Act
  - Lists the exclusive rights of a copyright holder (s.16(1)(a-e)) – Reproduction or Adaption
  - Extends these rights to the whole, or a ***substantial part*** of the protected work (s.16(3)(a))
- “There must be sufficient ***objective similarity*** between the infringing work and the copyright work, or a ***substantial part*** thereof, for the former to be properly described, not necessarily as identical with, but as a reproduction or adaptation of the latter.”

*Francis Day Hunter v Bron*

# Looks like Infringement?

## Sounds like infringement?

Austin v Columbia Gramophone

*“Infringement of copyright in music is not a question of note for note comparison, but whether the substance of the original copyright work is taken or not. It falls to be determined by the **ear as well as by the eye**”*

- Note-by-note comparison
- The auditory perception of musical similarity

# By the Eye



- Typically musical comparison 'by the eye'
  - Line Drawing
  - Highlighting
- Criticism
  - This approach has been criticised as 'simple', 'primitive', and 'misleading' (Cronin)
  - Invites a 'subjective and limited breakdown and analyses of songs [that] often lead to conflicting interpretations from experts' (Liebesman)

CHORUS 2nd time

1 2 3 4  
In a little Spanish town, 'Twas on a night like this,

5 6 7 8  
Stars were peek-a-boo-ing down, 'Twas on a night like this,

"WHY"

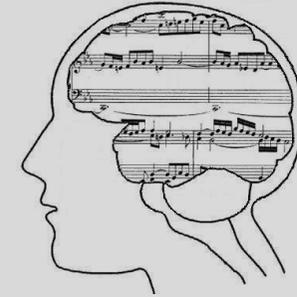
1 2 3 4  
I'll never let you go, WHY, be-cause I love you

5 6 7 8  
I'll al-ways love you so, WHY, be-cause you love me.

The image shows a musical score for the song 'By the Eye'. It consists of four systems of music, each with a vocal line and a piano accompaniment. The first system is labeled 'CHORUS 2nd time' and contains measures 1-4. The second system contains measures 5-8. The third system is labeled '"WHY"' and contains measures 1-4. The fourth system contains measures 5-8. Annotations include a green circle around the first measure of the first system, a green circle around the first measure of the third system, and a yellow highlight under the piano accompaniment in the second and fourth systems. Red arrows point from the green circles to the yellow highlights.

*Francis Day Hunter v Bron*

# By the Ear



- Auditory perception of similarity

Francis Day Hunter v Bron

*'Similar to the extent that an ordinary reasonably experienced listener might think that perhaps one had come from the other'*

*'The public has a purer approach to music than the critics.' That, of course, does not mean that one must discount the help that the critics can give, but I think I must rely on the ear as well as on the eye'*

- Williamson Music v Pearson and 'the reasonable listener survey'

# A Substantial Part

- CDPA s.16(3)(a) Extends copyright protection to the whole or a *substantial part* of the protected work
- What is a 'substantial part'?
  - Case-by-Case approach
  - The Point of reference

Designer Guild v Russell Williams

*"It depends upon its importance to the copyright work. It does **not** depend upon its importance to the defendants"*

- Quality over Quantity

Newspaper Licensing v Marks and Spencer

*"Quality should be identified; 'by reference to the reason why the work was given copyright protection'*

# A Substantial Part

- Idea vs. Expression of the ideas(s)

Designer Guild v Russell Williams

*Has the infringer incorporated a substantial part of the independent skill, labour etc. contributed by the original author in creating the copyright work?*

- Non protection for commonplace ideas

Designer Guild v Russell Williams

*'the more abstract and simple the copied idea, the less likely it is to constitute a substantial part'*

&

*'certain ideas expressed by a copyright work may not be protected because [. . .] they are not original, or so commonplace as not to form a substantial part of the work'*

- Musical's works

Creagh v. Hit and Run

*'[. . .] not original, forming as they do, notes 1, 2 and 3 of the minor scale and are commonplace'*

EMI v Papathanassiou

*'The [disputed part] was a musical commonplace and had been used by the defendant himself before the composition of "City of Violets"'*

# Things to Keep in Mind

- Similarity measurement in music is determined by the ear *and* the eye
- The evidence presented often uses third party music
- It is generally accepted that there is a presupposed level of knowledge from the listener
- Not every divisible part of a protected work is afforded copyright protection
- Whether or not a part constitutes a substantial part is always in reference to the protected work

# Studying music plagiarism empirically

Questions:

- How do court decisions relate to melodic similarity?
- Can they be predicted by similarity algorithms?
- How do listeners, algorithms, and courts agree?
- How important is modelling of prior musical knowledge?
- How to model plaintiff's vs defendant's perspective?

# The problem, e.g. *Selle v. Gibb* (1983, 567 F.Supp 1173)

- Ronald Selle, “Let It End”, 1975 (unpublished)



- Bee Gees, “How Deep Is Your Love” (1977)



# Two Studies

Müllensiefen & Pendzich (2009):

- 20 US cases on melodic plagiarism with binary decision (yes/no plagiarism)
- Different computational approaches (edit distance/string matching, n-grams, Tversky's ratio model of similarity)

Müllensiefen, Wolf, & Cason (in prep.):

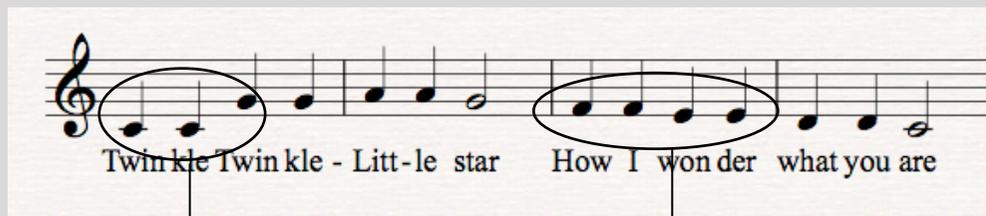
- 19 cases from US and Commonwealth (yes/no plagiarism)
- Different computational approaches (Tversky's ratio model, compression distance, Euclidean feature distance)
- 37 participants tested on implicit memory paradigm indicating similarity between tunes

# Measuring melodic similarity

1. Break melodies up into features
2. Weight features by commonness in pop music history
3. Compute similarity based on unique features shared between melodies

# 1) Breaking melodies up into features

Features: Short motives (m-types) similar to words in language



m-type of length 2:  
"s1e\_u5e"

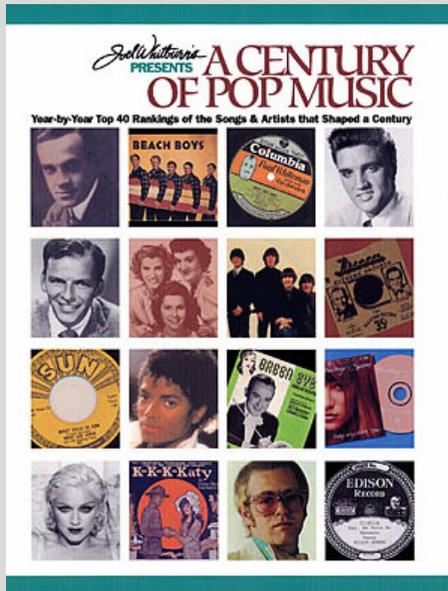
m-type of length 3:  
"s1e\_d2e\_s1e"

And then?

Count melody-types!

Word Type $\square$	Frequency $f(\square)$ ,	Melodic Type $\tau$ (pitch interval, length 2)	Frequency $f(\tau)$ ,
Twinkle	2	0, +7	1
little	1	+7, 0	1
star	1	0, +2	1
How	1	+2, 0	1
I	1	0, -2	3
wonder	1	-2, -2	1
what	1	-2, 0	2
you	1	0, -1	1
are	1	-1, 0	1

## 2. Weight features by commonness



- Count motives in Goldsmiths database (14,000 songs), representing popular music since 1950s
- Derive IDF weights (established from text retrieval)
  - Common motives: low weights
  - Rare and unique motives: high weights

### 3. Compute similarity: Tversky's ratio model (1977)

Rationale: Similarity of two objects,  $\sigma(s,t)$ , is related to

- Number of features  $s$  and  $t$  have common (vs. number of features they don't have in common)
- Perceptual salience of features,  $f()$
- Direction of comparison, often:  $\sigma(s,t) \neq \sigma(t,s)$

$$\sigma(s,t) = \frac{f(s_n \cap t_n)}{f(s_n \cap t_n) + \alpha f(s_n \setminus t_n) + \beta f(t_n \setminus s_n)}, \alpha, \beta \geq 0$$

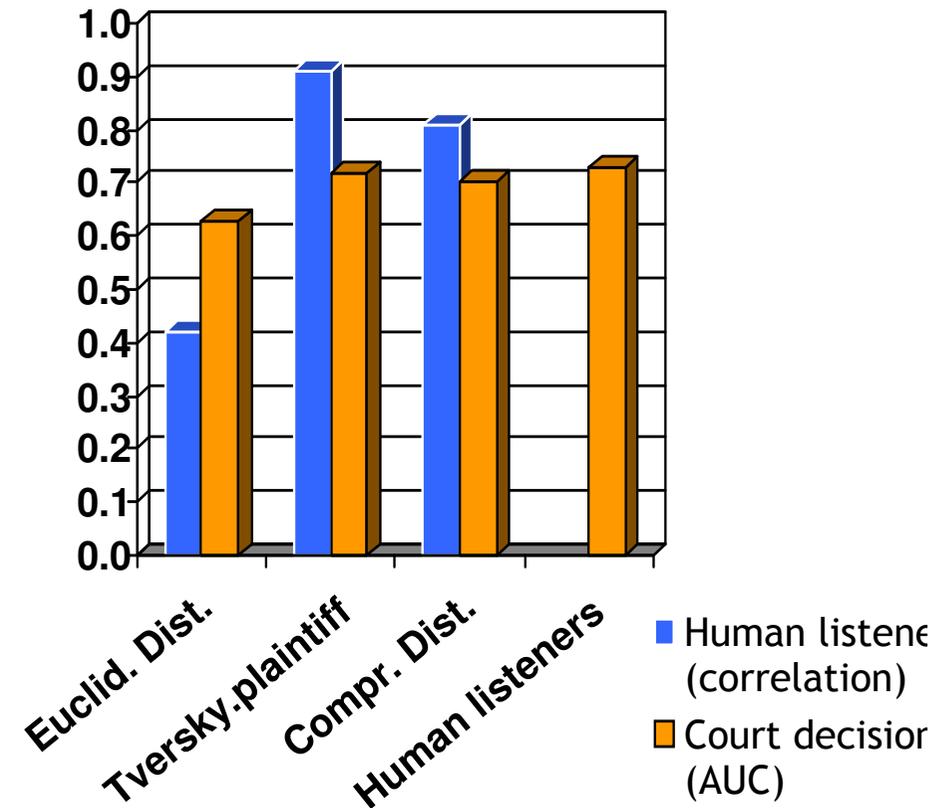
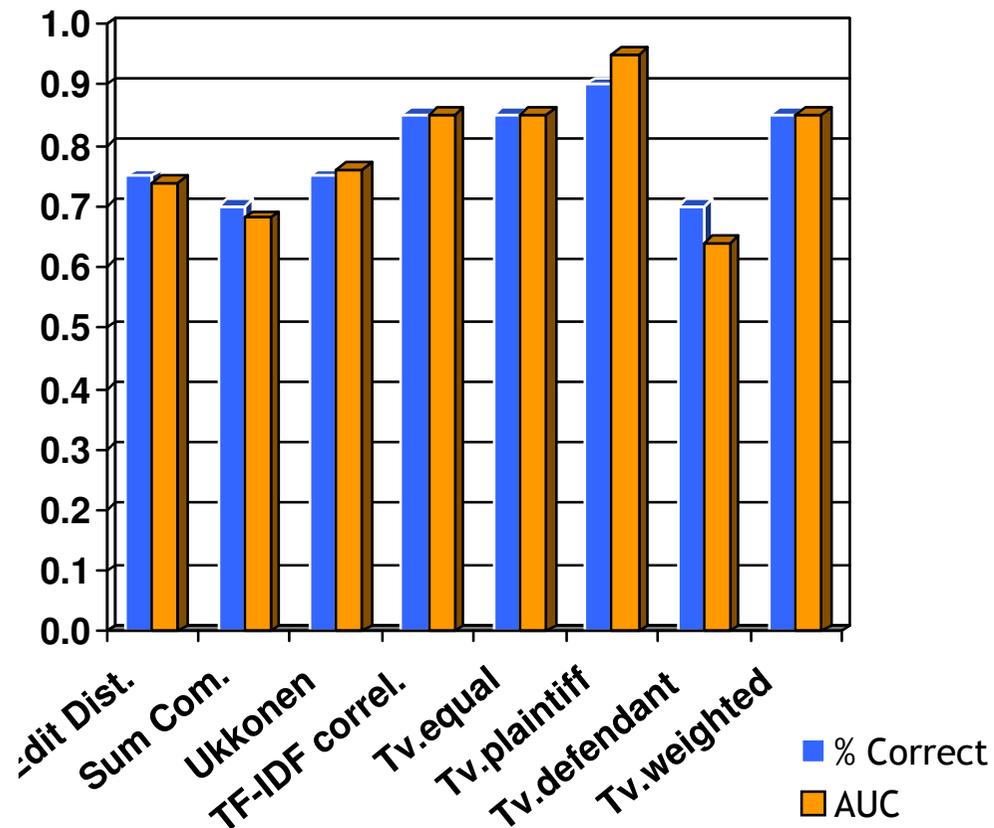
Implementation of ratio model for melodic similarity

- Objects => melodies
- Features => short motives
- Perceptual salience => IDF weights derived from pop database
- Different values of  $\alpha$ ,  $\beta$  to change frame of reference (plaintiff vs defendant)

# Empirical results

Müllensiefen & Pendzich, 2009

Müllensiefen, Wolf & Cason, in prep



Results:

- 1) Tversky's ratio model closest to court decisions and listener judgements
- 2) Absolute agreement comparable to group of 'reasonable listeners'
- 3) Modelling of plaintiff's perspective gives optimal results

# Tversky's ratio model - legal implications

The ratio model of similarity:

- Good empirical benchmarks
- Legally adequate?

$$\sigma(s,t) = \frac{f(s_n \cap t_n)}{f(s_n \cap t_n) + \alpha f(s_n \setminus t_n) + \beta f(t_n \setminus s_n)}, \alpha, \beta \geq 0$$

Implementation and legal interpretations of melodic similarity:

- *Objective Similarity*  $\Leftrightarrow$  Relative overlap in motives (numerical value)
- *Substantial Part*  $\Leftrightarrow$  Perceptual salience function
- *Non-protection of common place ideas*  $\Leftrightarrow$  Down-weighting of common elements
- *Knowledge of reasonably experienced listener*  $\Leftrightarrow$  Statistical information derived from pop corpus
- *Importance to copyright work not defendant's*  $\Leftrightarrow$  parameters  $\alpha, \beta$  to adjust for plaintiff's perspective

# To Conclude

- Tversky's ratio model can be implemented straightforwardly for measuring tune similarity
- Good agreement with court decisions and listeners' judgements
- Core components match key features of copyright act and case law
- Not subject-specific but based on general similarity perception
- Provides opportunity to interrogate legal concepts on empirical basis

## Open questions:

- Implementation of other musical elements (harmony, lyrics, sounds, polyphony)
- Applicable to continental author's right and legal practice?

# References

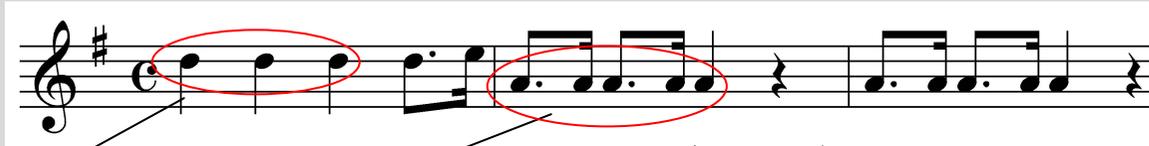
Müllensiefen, D., and Pendzich, M. (2009). Court decisions on music plagiarism and the predictive value of similarity algorithms. *Musicae Scientiae*, Discussion Forum 4B, 257-295.

Cason, R.J.S., & Müllensiefen, D., (2012). Singing from the same sheet: computational melodic similarity measurement and copyright law. *International Review of Law, Computers & Technology*, 26(1),25-36.

Müllensiefen, Wolf, & Cason, (in prep). Algorithmic models of human similarity perception and court decision on music plagiarism.

*Thanks very much for your attention!*

# Making Melodies Computable

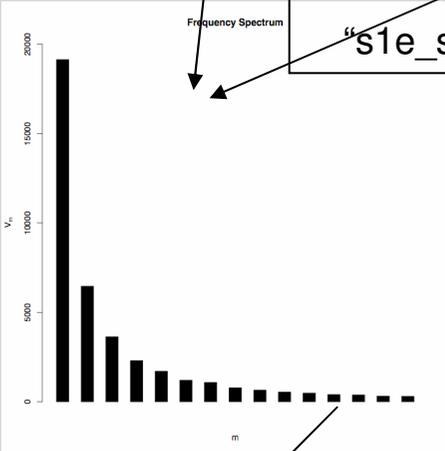


m-type of length 2:  
"s1e\_s1e"

m-type of length 4:  
"s1q\_s1l\_s1q\_s1l"

Symbol sequence encoding:

"s1e\_s1e\_s1q\_u2q\_d5l\_s1q\_s1l\_s1q\_s1l\_s1q\_s1q\_s1l\_s1q\_s1l"



Frequency Spectrum

$$i.abs.std = \sqrt{\frac{\sum_i (\Delta p_i - |\overline{\Delta p}|)^2}{N-1}} = 2.83$$

Overlap in m-types  
between s, t (Tversky)

Mutual  
compressability of s,t  
(Vitanyi)

Euclidean distance of  
global features between  
s,t (Shepard)