Is it the song and not the singer? Hit song science using structural features of melodies

Daniel Müllensiefen, Kelly Jakubowski
Goldsmiths, University of London
Klaus Frieler
Hochschule für Musik "Franz Liszt" Weimar

An old question:

Can a perfectly composed melody

make you go insane (Sirens, Ulysses)
heal (Pythagoras)
turn sorrows into joy (Magic Flute)
make walls tumble (Jericho)
make you will-less (Pied Piper of Hameln)
charm wild beasts and soften the heart of Death (Orpheus)

???

→ Great stories, often effective templates for media reports and scientific narratives (e.g. *the Mozart effect*)

More generally (and modestly)

Can musical structure influence / bias behaviour?

Can we predict human behaviour from musical structure?

Are all musical structures equally likely to generate certain cognitive or emotional responses? (H0)
or
Are certain structures more prone than others to trigger a certain reaction? (H1)

Empirical research since late 1960s, interdisciplinary area 'Experimental Aesthetics'.

Tune Features and ...

- **Singalong behaviour**: Vocal features that motivate people to sing along to music in nightclubs (Pawley & Müllensiefen, 2012).
- **Tune memory**: Melodic features that make melodies more memorable (Müllensiefen & Halpern, 2014).
- **The earworm formula**: Melodic features of songs that become earworms (Williamson & Müllensiefen, 2012; Müllensiefen, Jakubowski et al. in prep.).
- **The hitsong formula**: Prediction of commercial success of songs commonly based on audio features.

Hit Song Science (1)

- Dhanaraj & Logan (2005) classified hits vs. non-hits based on acoustic and lyrical features:
  - best acoustic classification rate = 0.66, slightly better for just lyrical features (0.68),
  - limitation- compared No. 1 hits to all other songs.

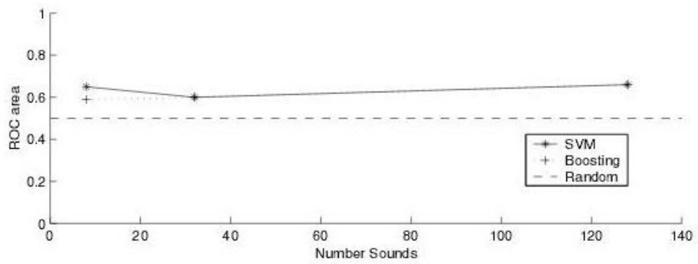


Figure 1: Average ROC area for acoustic-based features with various numbers of sounds for SVM and boosting classifiers

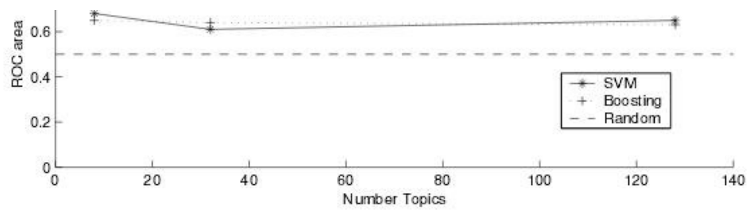


Figure 2: Average ROC area for lyric-based features with various numbers of topics for SVM and boosting classifiers

Hit Song Science (2)

- Ni et al. (2011) attempted to distinguish top 5 songs from songs in positions 30-40:
  - computed EchoNest features including tempo, time signature, song duration & loudness,
  - prediction accuracy around 0.57,
  - louder, harmonically simple, & faster songs preferred now.

http://scoreahit.com/

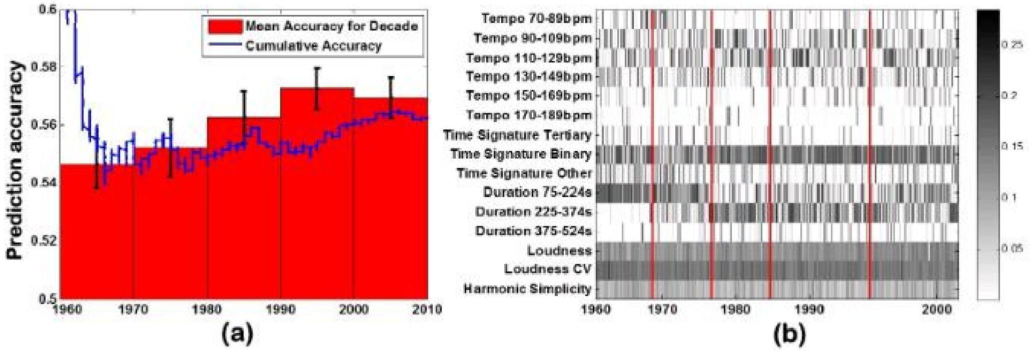
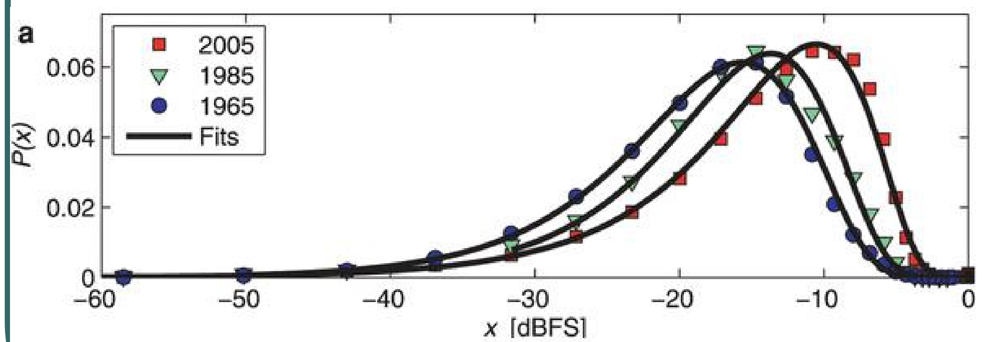
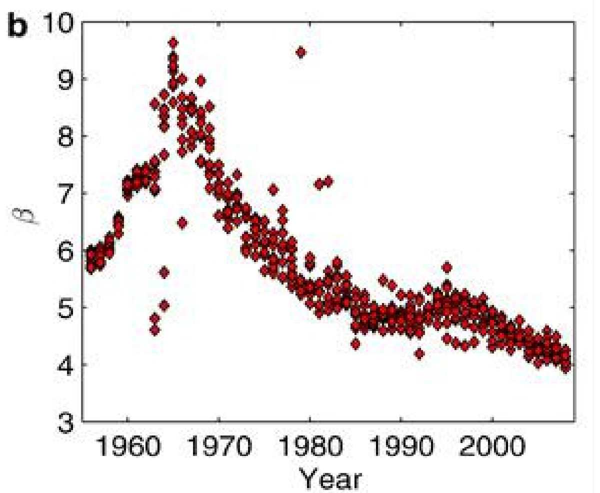


Figure 1: Prediction accuracy with standard deviation for decade (a) and evolution of classifier weights (b). Time progresses on the horizontal axes, decades are grouped for ease of interpretation.

Hit Song Science (3)

- Serra et al. (2012) looked at the evolution of hit songs over time on the basis of audio features.
  - Concluded music is getting louder, more timbrally homogeneous, and more restricted in terms of pitch patterns.





Hit Song Science (4)

- Nunes & Ordanini (2014) explored influence of instrumentation on commercial success:
  - hand-coded instrumentation data,
  - more popular songs always include backing vocals,
  - songs that contain an atypically low or high number of instruments tend to become hits.
  - No hit prediction accuracy.

Criticism of acoustic feature approach

- Pachet & Roy (2008) used a database of 32,000 songs and report that acoustic classifiers are not a good predictor of a song’s relative popularity (“Hit Song Science is Not Yet a Science”).

Alternatives to acoustic feature approach

- Social-knowledge driven hit song predictions using:
  - Early market responses to new songs (Salganik et al., 2006, Salganik & Watts, 2008),
  - Music social networks (Bischoff et al., 2009),
  - Twitter (Kim et al., 2014),
  - Analytics services (e.g. Musicmetric) used by Spotify, E and MTV to track fans and listening time via social media
- Melodic features
  - Kopiez & Müllensiefen (2011) predict commercial success (chart entry or not) of cover versions of songs from the Beatles' Revolver using two melodic features (pitch range and pitch entropy).
  - Accuracy of classification tree: 100%.
  - N=14

The melodic feature approach

1. Collect data on song popularity and commercial success (i.e., chart data).
2. Find transcription of corresponding tunes and extract melody features.
3. Data-mine melody features to explain commercial success.

Summary Features

The task: Summarise the content of a melodic phrase

 = ?

Cognitive Hypothesis: Listeners abstract summary representation of short melodies during listening

Format: Number that represents particular aspect of melody

Pitch range (p.range):

$$p.range = \max(p) - \min(p)$$

Melfeature & the MeloSpySuite

- MeloSpySuite: Tool box for computational analysis of melodies (melfeature, melconv, melpat).*
- melfeature uses a modular and extendable approach to extract arbitrary features from monophonic melodies.
- Core concept: Transformation chains.
- Scalar-, vector- and matrix-valued features.
- “Standard set” currently comprises several hundred of features.
- This study: Set of 152 scalar (summary) features.

*Available at http://jazzomat.hfm-weimar.de/download

Simple features: Categories

- Contour
  - Huron contour classes
  - Pitch extrema
- Interval:
  - Semitone interval
  - Interval classes
  - Parsons code
- Metre:
  - Metrical circle map
- Pitch
  - Pitches (raw, pc, tonal, chordal, chromatic/diatonic)
- Rhythm
  - Durations, IOIs
  - Event densities
  - nPVI, CV
- Sequence/Interval
  - Mean run-lengths (steps, thirds, chromatic etc.)
  - Bi- and Trigram distributions
  - Interval (classes), Parsons
- Sequence/Pitch
  - as above, using pitch representations
- Sequence/Rhythm
  - as above, using pitch representations
- Structure
  - Self-similarity
- Articulation & Auxiliaries
  - (not used here)

Simple features: Types

- Only intrinsic, no extrinsic, i.e., corpus-based features.
- Variation on a common theme: Descriptive statistics of distributions:
  - Mean, median, mode, std, var, min, max, range, etc
  - circular statistics where necessary,
  - single densities of feature classes,
  - entropy,
  - Zipf coefficients.
- Some specialised descriptors (e.g, Huron contour classes, nPVI, metrical complexities).

Similar Approaches

Folk Song Research / Ethnomusicology

- Bartók (1936), Bartók & Lord (1951)
- Lomax (1977)
- Steinbeck (1982)
- Jesser (1992)
- Sagrillo (1999)
- Volk et al. (2008)
- van Kranenburg, Volk, Wiering (2012)

Computational / Cognitive Musicology

- Eerola et al. (2001, 2007)
- McCay (2005)
- Huron (2006)
- Frieler (2008)
- Müllensiefen & Halpern (2014)

Popular Music Research

- Moore (2006)
- Kramarz (2006)
- Furnes (2006)

Computational Linguistics / Cognitive Psychology

- Baayen (2001)
- Landauer et al. (2007)
- Sedlmeier & Betsch (2002)
- Cortese et al. (2010)
- ….

Test case: Prediction of chart success

Do structural features make tunes more commercially successful?

The tune collection

- 266 songs taken from a related project on ‘features of earworm tunes’
  - half of these were tunes reported by participants as particularly catchy songs on the “Earwormery” database
  - the other half were songs selected specifically to be from similar artists and UK chart position
  - song, artist, chart data, and genre recorded from all songs
  - melody line from the part reported as most catchy by participants (or chorus) was extracted

The analysis

- Data:
  - Dependent variables: “No. weeks in charts”, “highest chart position”
    - ⇒ Binary classification into 122 “hits” vs 144 “non-hits” via k-means clustering (k=2)

Artist	Song	Genre	Sample
Queen	I'm Going Slightly Mad	Rock	
Queen	Bohemian Rhapsody	Rock	
Britney Spears	Circus	Pop	
Britney Spears	Toxic	Pop	
Lionel Richie	Running with the Night	Pop	
Lionel Richie	Hello	Pop Rock	



The analysis

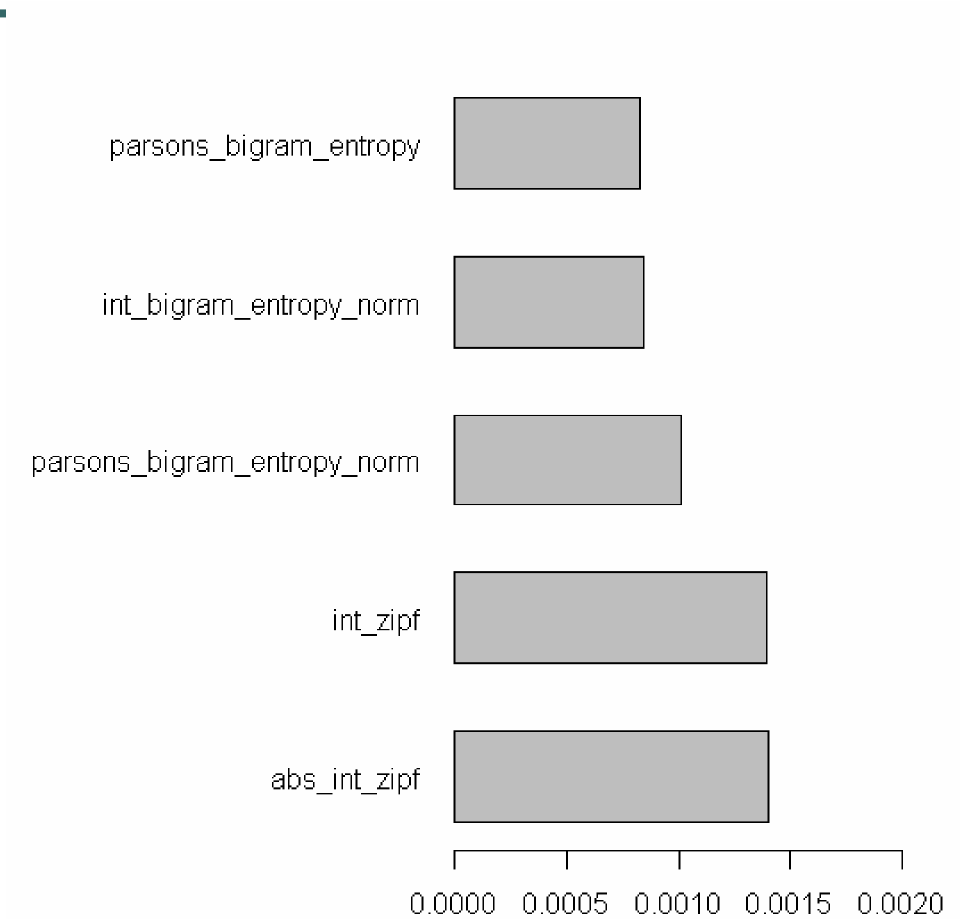
Artist	Song	Genre	Highest Entry	Weeks in Charts	Classification
Queen	I'm Going Slightly Mad	Rock	22	5	Non-hit
Queen	Bohemian Rhapsody	Rock	1	17	Hit
Britney Spears	Circus	Pop	13	18	Non-hit
Britney Spears	Toxic	Pop	1	14	Hit
Lionel Richie	Running with the Night	Pop	9	12	Non-hit
Lionel Richie	Hello	Pop Rock	1	15	Hit

The analysis

- Predictor variables: 150 melody features from MeloSpySuite
- Statistical model: random forest (Breiman, 2001)
  - Handles categorical and numerical data
  - Handles non-linear relationships
  - Handles ‘large k small n problem’
  - Implementation based on tree models using permutation tests (providing statistical framework, p-values, etc; Strobl et al., 2011)

The results: Random Forest

- (Honest) classification accuracy: ~ 52% (dishonest, i.e. overfitted accuracy: 92%)
- Most important features:
  - Zipf distribution of intervals
  - Entropy of Parsons code bigrams



The hit tune features

Normed entropy of Parsons code bigrams

Measures degree of uniformity of distribution of melodic movement (up, down, repetition bigrams).

The higher the entropy, the more variable the melodic movement, but not too much, since repetition of same movement must occurr as well.

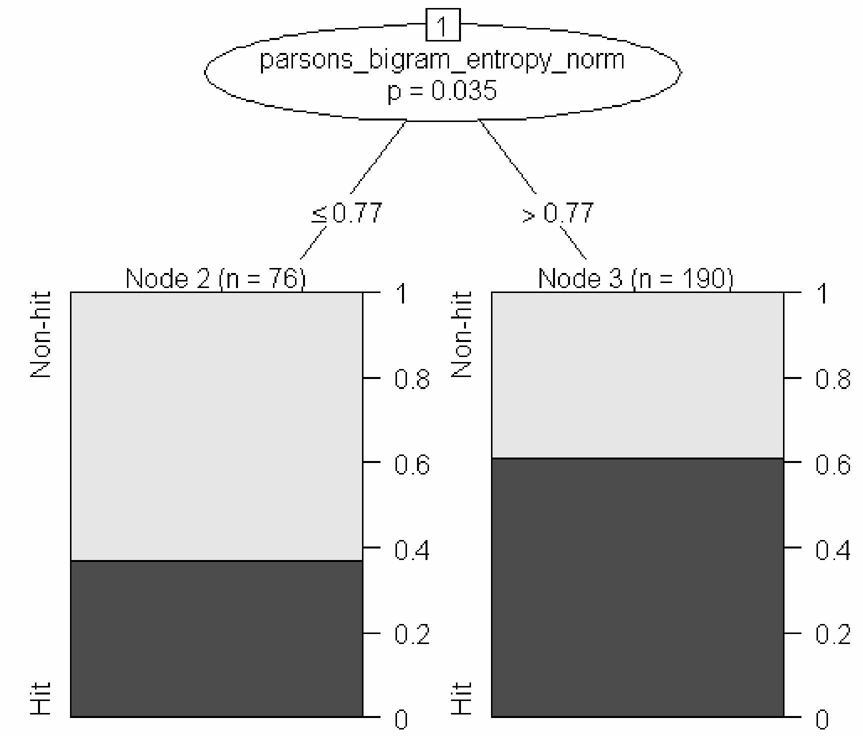
Zipf coefficient of intervals

Measures how dominant the most common intervals are. The higher the Zipf coefficient, the more the distribution is determined by a small set.

The results: Single tree model

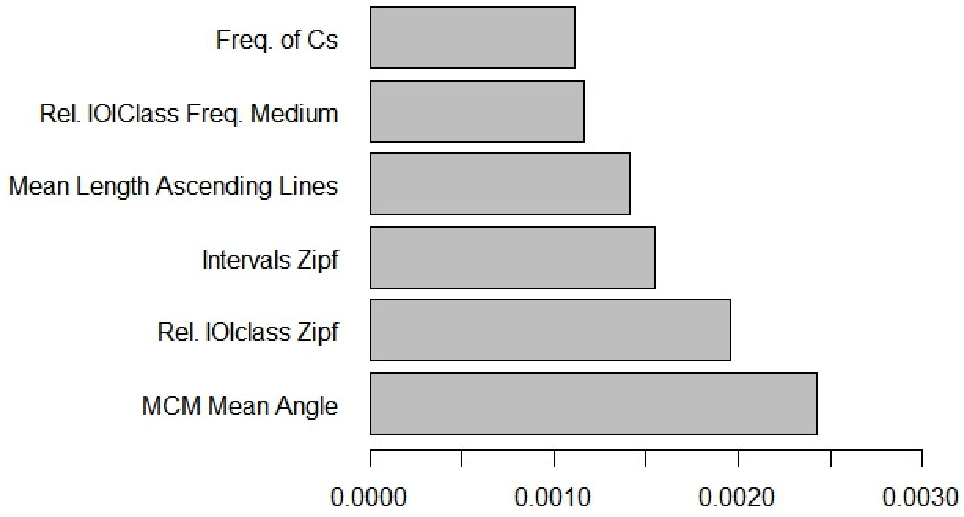
- Only significant feature:
  - Normed entropy of Parsons code bigrams

=> Tunes with more variables melodic contours are more commercially successful



The results: Earworm prediction

- (Honest) classification accuracy: ~ 55% (dishonest, i.e. overfitted accuracy: 98%)



The earworm features

Mean angle of Metrical Circle Map distribution

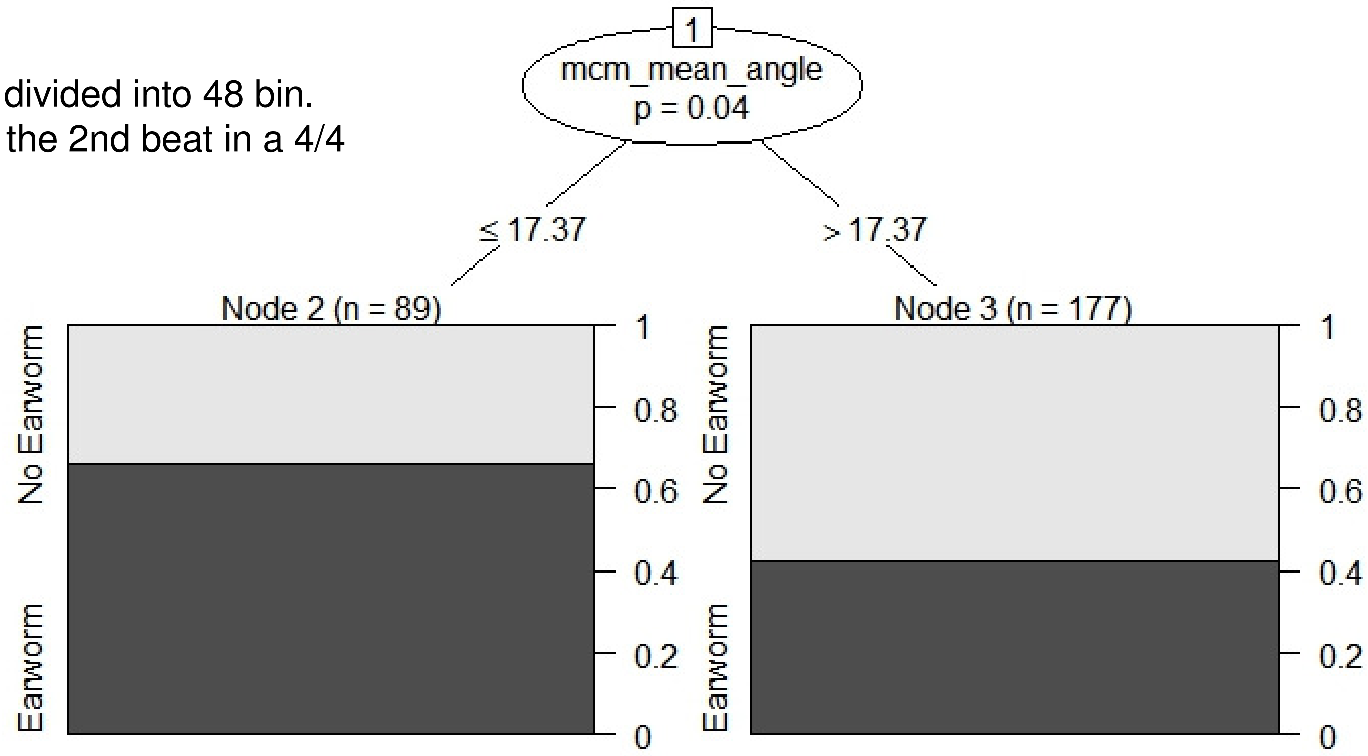
Measures the centroid of metrical positions across the melody,

Relative IOI classes Zipf coefficient

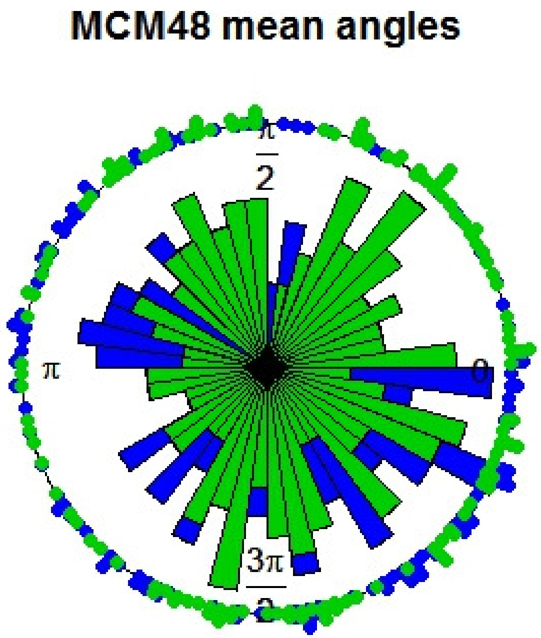
Measures how dominant the most common IOI classes are. The higher the Zipf coefficient, the more the distribution is determined by a small set of values.

The results: Earworm tree

Bars are divided into 48 bin.
Bin 16 is the 2nd beat in a 4/4



The results: Mean MCM angle



Green: Earworms

In summary

- Features of hit songs based on pitch contour and intervals
- Distribution of categorical feature values is important
- BUT: overall low explanatory (predictive) power

⇒ *Finding the secret formula of hit tune is at least as difficult as finding their magical combination of audio features!*

Next steps:

- Use distributional information from large pop corpus (2nd order features)
- Use pattern features (Melpat) as well corpus information
- Combine audio and melodic feature approach

The wider picture

Finding the secret formula of the perfect melody is difficult:

- Many potential features to consider => Aggregation / Feature selection
- People are very idiosyncratic in their responses to music and even pop music is a diverse terrain => different but equally effective formulae?
- Potentially many confounding factors

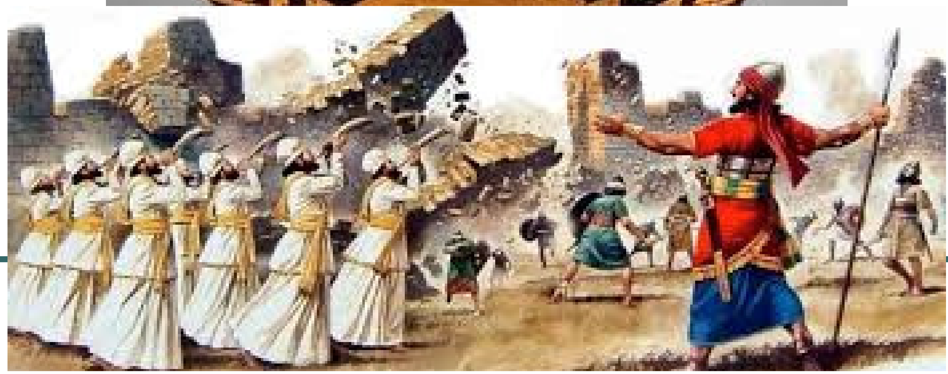
BUT:

- We found a few significant features of melodic structure
- Feature distribution within tune seems important

=> Points to general cognitive mechanism for event frequency processing

However

We’re still far from the golden times of music manipulation.



Is it the song and not the singer? Hit song science using structural features of melodies

Daniel Müllensiefen, Kelly Jakubowski
Goldsmiths, University of London
Klaus Frieler
Hochschule für Musik “Franz Liszt” Weimar