An empirical field study on sing-along behaviour in the North of England

Alisun R. Pawley
Department of Music, University of York

Daniel Müllensiefen
Department of Psychology, Goldsmiths, University of London
What do these songs have in common?

- I’m Always Here
- Livin’ on a Prayer
- Chelsea Dagger
Introduction

- Strong historical tradition of singing along in England
- 20th century technologies & professionalisation of singer suppress public singing
- Singing along in leisure contexts is one of few public music-making opportunities today
Past Research

- Social/Cultural Studies: **Social bonding, expression of identity, ‘neo-tribes’** (Maffesoli, 1988; Finnegan, 1989; Bennett, 1997; Björnberg and Stockfelt, 1996; Malbon, 1999; Jackson, 2004)

- Psychology: **Positive effects of vocalising** (Clift and Hancox, 2001; Freeman, 2001; Unwin, Kenny and Davis, 2002; Kreutz, et al, 2004; Clift, et al, 2007)

- Popular Music Analysis: ‘**Singable**’ melodies (Stefani, 1987), structural features of **popular anthems** (Dockwray, 2005)
Aims

- What motivates people to sing along to a song in a leisure context?
- Do songs have intrinsic features that make them ‘singalongable’?
Methods: Field Research

- Participant observer
- Quantitative & qualitative data
- 30 nights of research (Nov 2006 - Jul 2007)
- 5 venues: Manchester, Leeds, York & Kendal
- DJed & live music
Quantitative Results: The Data

- Dependent variable: percentage of people singing along
- Two sets of explanatory (predictor) variables: contextual & musical
  - 1050 ‘song events’
  - 636 different songs
  - 332 song events used in musical analysis (121 songs)

- Contextual variables:
  - Place of song in set
  - Day of week
  - Venue size & function
  - Live vs recorded
  - Age range of audience
  - Date of release, UK chart position, weeks in UK chart

- Musical variables (34 total):
  - Vocal span & phrase lengths
  - Vocal hook
  - Vocal performance
  - Lyrics
  - Gender
  - …
Distribution of Sing-along Percentages (n=1050)
**Songs with highest sing-along percentages akin to Dockwray’s (2005) Rock Anthems**

**‘Top Ten’ Sing-Along Songs Observed Twice or More**

<table>
<thead>
<tr>
<th>Song</th>
<th>Average % of people singing along</th>
<th>Average no. in audience when song was played</th>
<th>No. of song events featuring song</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>We are the Champions</em> (Queen)</td>
<td>85.91</td>
<td>85</td>
<td>4</td>
</tr>
<tr>
<td><em>Y.M.C.A.</em> (Village People)</td>
<td>85</td>
<td>75</td>
<td>2</td>
</tr>
<tr>
<td><em>Fat Lip</em> (Sum41)</td>
<td>81.58</td>
<td>80</td>
<td>3</td>
</tr>
<tr>
<td><em>The Final Countdown</em> (Europe)</td>
<td>78.89</td>
<td>70</td>
<td>2</td>
</tr>
<tr>
<td><em>Monster</em> (The Automatic)</td>
<td>78.72</td>
<td>70</td>
<td>4</td>
</tr>
<tr>
<td><em>Ruby</em> (Kaiser Chiefs)</td>
<td>78.52</td>
<td>68.33</td>
<td>3</td>
</tr>
<tr>
<td><em>I’m Always Here</em> (Jimi Jameson)</td>
<td>77.18</td>
<td>82.86</td>
<td>7</td>
</tr>
<tr>
<td><em>Brown Eyed Girl</em> (Van Morrison)</td>
<td>76.85</td>
<td>86.67</td>
<td>3</td>
</tr>
<tr>
<td><em>Teenage Dirtbag</em> (Wheatus)</td>
<td>75.7</td>
<td>68</td>
<td>5</td>
</tr>
<tr>
<td><em>Livin’ on a Prayer</em> (Bon Jovi)</td>
<td>75.36</td>
<td>87</td>
<td>7</td>
</tr>
</tbody>
</table>
Analysis 1: Contextual Variables

- Who sings along at what time and in which context?
- Use statistical regression tree to cope with ‘messy’ data:
  - Outliers
  - Non-linear relationships
  - Violated assumptions of normality and variance homogeneity
Tree Model: Contextual Variables

Conditional Inference Regression Tree model: explains ~40% of variance in the data
Analysis 2: Musical Variables

- Which **musical features** motivate sing-along behaviour?
- Regression tree with musical variables: $R^2 \sim .05$
  - *Maybe no single sing-along formula?*
  - *Try different subsets and complex interactions of musical features*

- Random forest (Breiman, 2001; Strobl et al., 2009) regression:
  - Build (‘grow’) many tree models on data subsets each with a subset of the explanatory variables
  - Use majority vote of trees in forest to decide on prediction value for each case
  - *Pro: Much better prediction accuracy than from single tree*
  - *Con: No simple rules or individual graphical model but variable importance index*
Results

Prediction accuracy:
- All variables: $R^2 = .65$
- Contextual variables only: $R^2 = .40$
- Musical variables only: $R^2 = .52$

Most important variables (importance index):
1. Combined model from contextual variables (101.4)
2. High chest voice (6.8)
3. Vocal effort (6.4)
4. Gender of vocalist (4.9)
5. Clarity of consonants (3.4)
6. Vocal melisma and embellishment (3.4)

Then:
- Compositional structure of melody
- Features of lyrics
- ...

Aspects of vocal performance
Relating most important predictors to singing percentages (by single trees)

Low effort

High effort

No use of high chest voice

Use of high chest voice
Very clear consonants

Less clear consonants

More vocal embellishments

Less vocal embellishments
Summary

- Contextual and musical factors determine how many people sing along in leisure contexts (explained variance: ~65%)

- Singing along is positively affected by these contextual factors:
  - Larger venues
  - Younger people
  - Weekends
  - Songs played later in the set
  - Songs that spent 4 or more weeks in the charts

- Singing along is positively influenced mainly by factors relating vocal performance.
  => It’s the singer not the song!
Implications and Interpretations

- Contextual variables that encourage singing along can be connected with general revelry
- Familiarity & popularity potentially linked to singing along
- No single ‘sing-along’ formula for music
- Musical factors that do influence singing along are similar to qualities of anthems in popular music (Dockwray, 2005):
  - ‘Call to party’ – ‘tribal’ bonding
  - Expression of excitement of revellers
  - Word clarity: ease of understanding & reproduction
  - Singer expresses qualities that inspire confidence
  - Men don’t like to sing along to women
Next Steps

- Validate importance of musical variables for
  - songs that are currently popular
  - different social and cultural setting
  - different musical repertoire
- Extract musical features using computational tools
- Explore practical applications …?

*We are very open for collaborations!*
What we didn’t find ...